Psychometrika

CONTENTS

VOLUME ONE • JUNE 1936 NUMBER TWO

PSYCHOMETRIKA—VOL. 1, NO. 2
JUNE, 1936

SOME PROPERTIES OF THE COMMUNALITY IN MULTIPLE FACTOR THEORY

MERRILL ROFF*

Indiana University, Bloomington, Indiana

*The author wishes to express his appreciation of the encouragement and assistance given him by Dr. L. L. Thurstone.

Several theorems concerning properties of the communaltiy of a test in the Thurstone multiple factor theory are established. The following theorems are applicable to a battery of n tests which are describable in terms of r common factors, with orthogonal reference vectors.

1. The communality of a test j is equal to the square of the multiple correlation of test j with the r reference vectors.

2. The communality of a test j is equal to the square of the multiple correlation of test j with the r reference vectors and the n—1 remaining tests.

Corollary: The square of the multiple correlation of a test j with the n—1 remaining tests is equal to or less than the communality of test j. It cannot exceed the communality.

3. The square of the multiple correlation of a test j with the n—1 remaining tests equals the communality of test j if the group of tests contains r statistically independent tests each with a communality of unity.

4. With correlation coefficients corrected for attenuation, when the number of tests increases indefinitely while the rank of the correlational matrix remains unchanged, the communality of a test j equals the square of the multiple correlation of test j with the n—1 remaining tests.

5. With raw correlation coefficients, it is shown in a special case that the square of the multiple correlation of a test j with the n—1 remaining tests approaches the communality of test j as a limit when the number of tests increases indefinitely while the rank of correlational matrix remains the same. This has not yet been proved for the general case.

In the multiple factor theory of Professor L. L. Thurstone (2) the concept of communality has a prominent place. The communality of a test is defined as its common factor variance, or "that part of its variance which is due to factors common to other tests in the battery" (2, pp. 62-63). When standard scores are used, the total variance of a test can be expressed as follows (2, pp. 55,57):

$$\sum_{m=1}^{r} a_{jm}^2 + b_{jj}^2 + c_{jj}^2 = 1 ,$$

in which

m refers to the common factors,

j refers to tests,

r = the number of common factors,

$a_{jm} \equiv$ the loading of the common factor m in test j,

$b_{jj} \equiv$ the loading of the specific factor of test j in test j,

$c_{jj} \equiv$ the loading of the error factor of test j in test j.

The following terms have also been defined **(2, p. 63)**:

$$\sum_{m=1}^{r} a_{jm}^2 \equiv h_j^2 \equiv \text{communality of test } j \ ,$$

$$b_{jj}^2 + c_{jj}^2 \equiv u_j^2 \equiv \text{uniqueness of test } j \ ,$$

so that

$$h_j^2 + u_j^2 = 1 \ .$$

In factoring a correlation matrix it is desirable to use the communality, h_j^2, of each test as the diagonal entry for that test in the reduced correlational matrix. However, h_j^2 can be determined only after the factorization has been completed. Consequently, it is necessary to estimate the communality before the analysis begins, and several methods of estimation have been discussed by Thurstone **(2, pp. 85-91)**. Beyond this practical problem arises the theoretical problem of the characteristics and significance of this important concept, which is not yet completely solved.

The present paper describes several properties of the communality, and relates the factor methods to the older technique of multiple correlation. A new method of estimating the communality of a test is indicated, but whether or not this method will prove better than those now in use remains to be seen. The following theorems are applicable to a battery of n tests which are describable in terms of r common factors, with orthogonal reference vectors. Since the communality of all tests in a battery must remain invariant under rotation, the theorems which follow are quite independent of any specific position of the reference vectors.

1. The communality of a test j is equal to the square of the multiple correlation of test j with the r orthogonal reference vectors. That is,

$$h_j^2 = r^2_{j \cdot 123 \cdots r}$$

where

$r^2_{j \cdot 123 \cdots r} =$ the square of the multiple correlation between a test j and the r reference vectors.

Here when r is used as a correlation coefficient, it will always be accompanied by subscripts; when r occurs without subscripts, it will always refer to the rank of the matrix.

The loading of the common factor m in test j, a_{jm}, may be represented geometrically as the projection of the vector of test j on the

reference vector of m; it may also be regarded as the correlation between test j and the reference vector of m; then

$$a_{jm} = r_{jm} \; .$$

The multiple correlation coefficient, $r_{j.123\cdots r}$, is given by the well-known formula,

$$r_{j\cdot 123\cdots r} = \sqrt{1 - \sigma^2_{j\cdot 123\cdots r}}$$

where

$$\sigma^2_{j\cdot 123\cdots r} = \sigma_{j1}{}^2 \cdot \sigma_{j2\cdot 1}{}^2 \cdots \sigma^2_{jr\cdot 123\cdots (r-1)}$$

$$= (1 - r_{j1}{}^2)(1 - r_{j2\cdot 1}{}^2) \cdots (1 - r^2_{jr\cdot 123\cdots (r-1)})$$

where σ with appropriate subscripts denotes the standard error of estimate with multiple or partial correlation coefficients, the subscript j refers to the test, and the subscripts 1, 2, 3, $\cdots$ r refer to the r orthogonal reference vectors. It is interesting to observe the behavior of the partial coefficient under these conditions. The partial coefficient $r_{j2\cdot 1}$ is given by the formula

$$r_{j2\cdot 1} = \frac{r_{j2} - r_{j1} \cdot r_{12}}{\sqrt{(1 - r_{j1}{}^2)(1 - r_{12}{}^2)}} \; .$$

Since r_{12} is the correlation between two orthogonal reference vectors, it equals zero, and we have

$$r_{j2\cdot 1} = \frac{r_{j2}}{\sqrt{1 - r_{j1}{}^2}} = \frac{r_{j2}}{\sigma_{j1}} \; .$$

Thus

$$\sigma_{j2\cdot 1}{}^2 = 1 - \frac{r_{j2}{}^2}{\sigma_{j1}{}^2} = \frac{\sigma_{j1}{}^2 - r_{j2}{}^2}{\sigma_{j1}{}^2}$$

and

$$\sigma_{j\cdot 12}{}^2 = \sigma_{j1}{}^2 \cdot \frac{\sigma_{j1}{}^2 - r_{j2}{}^2}{\sigma_{j1}{}^2} = \sigma_{j1}{}^2 - r_{j2}{}^2 \; .$$

Continuing the process with an additional test,

$$\sigma_{j\cdot 123}{}^2 = \sigma_{j\cdot 12}{}^2 \cdot \sigma_{j3\cdot 12}{}^2 \; .$$

The coefficient $r_{j3\cdot 12}$, and all additional partial coefficients of higher order, will simplify in the manner shown above, that is, since all the variables exclusive of j are uncorrelated, the second term in the numerator and the second term in the denominator both vanish, so that

$$r_{j3\cdot 12} = \frac{r_{j3}}{\sigma_{j\cdot 12}}$$

and

$$\sigma_{j3\cdot12}{}^2 = \frac{\sigma_{j\cdot12}{}^2 - r_{j3}{}^2}{\sigma_{j\cdot12}{}^2} .$$

Then

$$\sigma_{j\cdot123}{}^2 = \sigma_{j\cdot12}{}^2 \cdot \frac{\sigma_{j\cdot12}{}^2 - r_{j3}{}^2}{\sigma_{j\cdot12}{}^2} = \sigma_{j\cdot12}{}^2 - r_{j3}{}^2$$

$$= 1 - r_{j1}{}^2 - r_{j2}{}^2 - r_{j3}{}^2 .$$

This process can be continued until the rth factor is reached. Since all the residuals vanish when the rth factor has been removed, all the remaining entries in the matrix must be zero, and neither the communality nor the multiple correlation coefficient can be increased by adding further variables. Thus

$$\sigma^2{}_{j\cdot123\ldots r} = 1 - r_{j1}{}^2 - r_{j2}{}^2 - \cdots - r_{jr}{}^2 ,$$

and $r_{j(r+1)}{}^2$ and all succeeding terms equal zero.
But $r_{jm} = a_{jm}$, consequently

$$\sigma^2{}_{j\cdot123\ldots r} = 1 - \ (a_{j1}{}^2 + a_{j2}{}^2 + \cdots + a_{jr}{}^2) \ ;$$

also

$$h_j{}^2 = a_{j1}{}^2 + a_{j2}{}^2 + \cdots + a_{jr}{}^2 .$$

Hence,

$$\sigma_{j\cdot123\ldots r}{}^2 = 1 - h_j{}^2 \ ;$$

$$r^2{}_{j\cdot123\ldots r} = 1 - \sigma^2{}_{j\cdot123\ldots r} .$$

Therefore,

$$h_j{}^2 = r^2{}_{j\cdot123\ldots r} .$$

Since the correlation between any two tests j and k is zero after the r common factors have been removed, a second theorem follows immediately.

2. The communality of a test j is equal to the square of the multiple correlation of test j with the r reference vectors and the n—1 remaining tests.

From this follows a corollary which will be used later.

Corollary: The square of the multiple correlation of a test j with the n—1 remaining tests is equal to or less than the communality of test j. It cannot exceed the communality.

This is readily seen to be so from the fact that $r_{j\cdot123\ldots r}{}^2$ cannot be increased by the addition of any or all of the tests of the battery.

The question at once arises whether the communality is equivalent to the square of the multiple correlation of a test j with the n—1 remaining tests without the reference vectors. Each reference vector

may be regarded as a variable with a communality of unity. When raw correlation coefficients are used, no test can have a communality of unity because of the presence of an error factor. As the number of tests is increased while the rank of their correlation matrix remains unchanged, the communality of any test must remain constant, while it can be shown that in general, under the same circumstances, the multiple correlation coefficient may increase. Consequently the two need not be equal, and the precise relationship is not determined. However, another theorem can be stated at this point.

3. The square of the multiple correlation of a test j with the n—1 remaining tests equals the communality of test j if the group of tests contains r statistically independent tests each with a communality of unity.

When these conditions are met, the r reference vectors can be rotated to coincide with the r statistically independent tests, and the above proofs will apply. These conditions can never be exactly realized with raw correlation coefficients because of the error factors.

When correlation coefficients which have been corrected for attenuation are used, it is possible for the communality of a test to approach or reach unity. If the number of tests in a battery increases indefinitely while the rank of the correlational matrix, and thus the communality, remains the same, it is then possible to find r statistically independent tests with a communality of unity and then the square of the multiple correlation of a test j with the n—1 remaining tests equals its communality. Thus:

4. With correlation coefficients corrected for attenuation, when the number of tests increases indefinitely while the rank of the correlational matrix remains constant, the communality of a test j equals the square of the multiple correlation of test j with the n—1 remaining tests.

With uncorrected coefficients the communality of a test must remain constant when new variables are added, if the rank of the correlational matrix remains the same, while the multiple correlation coefficient generally increases under these conditions. However, it was shown above that the square of the multiple correlation of a test j with the n—1 remaining tests cannot exceed the communality of the test. This suggests that the square of the multiple correlation of a test j with the other tests might approach the communality as a a limit when the number of tests increases indefinitely while the rank of the correlational matrix remains the same. This can be shown to be so in a special case as follows. If all the coefficients r_{j1} in a column of a correlational matrix are equal, and if all the other coefficients

r_{jk} are also equal, although not necessarily equal to r_{j1}, the formula for the multiple correlation becomes **(1, p. 311)**

$$r_{1 \cdot 234 \cdots n} = r_{j1}\sqrt{\frac{n}{1+(n-1)r_{jk}}} .$$

As n increases indefinitely,

$$\lim_{n\to\infty} r_{1.234 \cdots n} = \frac{r_{j1}}{\sqrt{r_{jk}}},$$

and

$$\lim_{n\to\infty} r^2_{1 \cdot 234 \cdots n} = \frac{r_{j1}^2}{r_{jk}} .$$

Under these conditions the correlational matrix is of rank one, and the communality of any test can be determined as

$$h_1^2 = \frac{r_{j1} \cdot r_{j1}}{r_{jk}} = \frac{r_{j1}^2}{r_{jk}} .$$

Thus,

$$\lim_{n\to\infty} r^2_{1 \cdot 234 \cdots n} = h_1^2 .$$

It seems likely that this conclusion is true in the general case, but this has not yet been proved.

The square of the multiple correlation of a test with the remaining tests in a battery can safely be used as an estimate of the communality of a test for practical purposes, for it cannot exceed the communality and will generally give a fair approximation. Whether this procedure gives more satisfactory results than methods now in use is a matter of fact which is yet to be determined. Since the multiple correlation process becomes very unwieldy as the number of variables increases, it is desirable to have a method of approximating the value wanted. It is possible to select a small number of tests from a battery which will give almost all the information for this particular purpose which could be obtained from a complete battery. A method of determining which variables to select for use here has been developed and will appear later.

REFERENCES

1. Holzinger, K. J. Statistical methods for students in education. New York: Ginn & Co. Pp. vii + 372.
2. Thurstone, L. L. The vectors of the mind. Chicago: University of Chicago Press. Pp. xv + 266.

PSYCHOMETRIKA—VOL. 1, NO. 2
JUNE, 1936

ON THE USE OF MATHEMATICS IN PSYCHOLOGICAL THEORY
(Concluded from previous issue)

J. F. BROWN
University of Kansas

IV. NON-METRICIZED DYNAMICAL VARIANTS OF THE PSYCHOLOGICAL FIELD

The topological concepts allow us to assign individuals and goals to certain spatial regions. They allow us to designate what locomotions are possible for an individual and what regions must be traversed in attaining a definite goal. But the topological concepts alone tell us nothing of the actual locomotions performed in psychological activities. So far we have seen that the position of the individual at the start of the psychological activity may be defined in reference to this. Psychological activities may be ordered to *locomotions.* The individual in this sense has the character of a *thing.* The space through which the locomotion occurs has the properties of a *medium.* (Cf. Heider (10). In physical problems where the field construct is used, one may make this same distinction. Bodies falling in the earth's gravitational field have the properties of things, while the atmosphere is to be characterized as a medium. Similarly in the electro-static field, the isolated conductors have the properties of things and the field has those of a medium. In the following, individual point-regions will be considered as things, and the fields in which the locomotions occur as media. Media have dynamic properties such as fluidity, permeability, cohesiveness and the like.* Our next step must be to introduce these concepts, define them, and give example of their use. These will only be illustrations; references to exact analyses in terms of a field theory will be given in Section VI. Such concepts are only of definite scientific value when they are capable of *operational definition.* In other words, assignment of a certain fluidity to a field is only permissible when it is done on a definite experiential or experimental basis. They may perhaps be useful for speculation, before operational definition is possible. Our language of data is rich in phenomenological descriptions where non-metricized dynamic concepts are used and readily comprehended, and the adoption of them as constructs undoubtedly originates in such phenomenological descriptions. But *sci-*

*Such dynamic terms have been frequently used by some sociologists and psychologists, usually without precise definition and without the realization that they represent theoretical constructs. It is hoped that the field-theoretical concepts will not be confused with these others.

entific meaning accrues to such concepts only when they may be precisely, i.e., operationally, defined. At the present time we may define these concepts in terms of *experimental* or *statistical indices.* We use the term index purposely to distinguish such numerical assignations from real or fundamental measurement. Experimental indices are gained from actual experiments and have their greatest use in problems of individual psychology. Statistical indices are used chiefly in problems of social psychology and sociology. Thus it is possible to use income-tax returns as an index for the permeability of boundaries separating social classes, or from questionnaire results to obtain indices for the variation of field fluidities amongst social groups.

In the following we shall suggest different operations which may be used in defining each concept. We do *not* wish at this early point in the development of field theory to commit ourselves to too definite procedures in defining our concepts, because, where several indices are available, it may later transpire that one of these has a particular value for a final definition. *In each specific problem where these concepts are used, it is necessary to give them an operational definition.* Unless this is done, the theory may become meaningless.

The non-metricized dynamical concepts which we will use are fluidity, degree of freedom, of social locomotion, permeability, tension, and vector. The application of such concepts to psychological fields where the psychological space is quasi-physical, i.e., where initial position and goal may be ordered to infinitely structured space (problems of mazes, circuitous routes, etc.), is quite obvious. For instance the permeability of an electrical grid as a barrier may be assigned an index figure on the basis of strength of electric shock. The strength of a vector towards the goal of a maze may be nonmetrically indicated by hours of hunger. For this reason we will deal chiefly with examples of problems where initial position and goal are not so easily defined.

Fluidity. By the degree of fluidity of a medium is meant the ease of locomotion in the medium.* Ease of locomotion depends not only on the fluidity of the medium, but also on the distribution of barriers in the medium and on internal psychological factors. It has meaning, however, to speak of the varying fluidity of psychological fields in themselves. For cases of actual physical locomotion it is quite obvious that *ceteris paribus* locomotion by walking across a street is in a more fluid medium than swimming a stream of equal breadth. We speak, however, of the fluidity of psychological fields which have no imme-

*Lewin uses fluidity in a somewhat different sense and believes that only under special conditions may the ease of locomotion be used as a criterion for fluidity.

diate physical correlate, and of the fluidity of social fields. Phenomenally one "moves about" more easily in daydreaming than in perceiving. Day dreams normally occur in a plane of lesser reality than perception and one is justified in assigning a greater fluidity to fields of lesser reality than to fields of greater reality.* Under such conditions fluidity may be operationally defined through the rate of diffuse discharge of tensions in the different fields. The memory for perceived acts and phantasied acts may be used for gaining an index-figure to designate the fluidity. If the tensions in both fields may be considered equal, then perception may be said to occur in a field of less fluidity than phantasy when more perceived acts are remembered than phantasied acts. It goes without saying that such experiments require regular serial variation, control of motivation and the other usual psychological controls.

Likewise social fields may be said to vary in fiuidity where fluidity means the ease of social locomotion. One speaks popularly of "stiff" formal parties and compares these with "free" Bohemian ones. The formal party is to be ordered to a field of low fluidity, the Bohemian party is to be ordered to one of high fluidity.

Degree of freedom of social locomotion. By degree of freedom of social locomotion is meant the comparative number of directions in which social locomotion is possible. In a field having a high degree of freedom of social locomotion many locomotions are possible compared with a field having a low degree of such freedom. In general the degree of freedom of social locomotion varies inversely with the number of barriers within the field. The various social classes are to be ordered to fields of varying degrees of freedom of social locomotion. The bourgeoisie is to be ordered to a field of high degree of freedom of social locomotion, the petite bourgeoisie to a field of medium degree of freedom, and the proletariat to a field of low degree. Index figures may be assigned to the degree of freedom of such fields on the basis of economic and sociological statistics regarding income, consumption, education, and the like. The various differences in degree of freedom of social locomotion is indicated in Figure VI. There is a close coordination between the number of barriers and their permeability of which we will next speak.

Permeability. By the degree of permeability of a barrier is meant the ease with which locomotions are executed through the barrier. Here one distinguishes between group- and inner-barriers. (Cf. above.)

*Cf. the experiments of Brown, of Dembo, and of Mahler, reported in Lewin (16).

The group-barrier of the Catholic Church may be said to be less permeable than that of Protestant denominations. One can join most Protestant sects by simply going to the meetings, whereas to obtain membership-character in the Catholic region it is necessary to take instruction, become baptized, etc. Operationally then we are quite

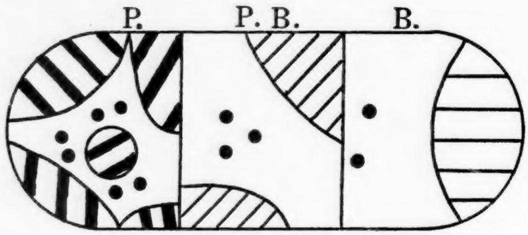


FIGURE VI

justified in saying that the barrier permeability of the Catholic Church region is less than that of the Protestant. Figure VII gives the dy-

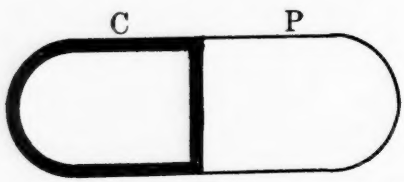


FIGURE VII

namical characterization of this situation. Differences in permeability will be shown by thickness of boundary. Similarly one may speak of differences in barrier permeability for other groups and define the concept operationally. The boundary separating nations might be assigned an index of permeability on the basis of immigration statistics, that separating class groups within a nation on the basis of income statistics, etc.

Inner-barriers, which represent impediments to locomotion within social field regions likewise vary in their permeability. The barriers to which laws are ordered are more permeable in the field of the bourgeois than in the field of the proletariat. Operationally this may be indicated by the ease with which bail and council are obtained by the bourgeoisie in comparison with the proletariat. Similarly, such taboos as being late to work represent barriers of decidedly different

permeability for the executive, the salaried worker, and the wage earner.

Vectors. The forces activating all locomotions in the psychological field are to be ordered to the concept of vector. These vectors represent forces causing psychological locomotion and are *directed* magnitudes. Their analogies in physical fields are the lines of field force within these fields. Such vectors, which represent forces, are to be indicated by arrows whose direction indicates the direction of the force, whose length represents its magnitude and whose point of application is at the point of the arrow. (Cf. Figure I.) Vectors are also used to indicate locomotions as in Figures II, III, etc. Hence vectors represent the psychological force concepts. We say that the magnitude of vector varies directly with the ease of locomotion through fields and barriers of constant fluidity and permeability.

In all cases of actual physical locomotion vectors may be assigned index figures (though *not* measured) on the basis of hours of hunger, the strength of electric shock which will be suffered in attaining a definite goal, etc. Such procedures are so well known to experimental psychologists that further elucidation of them is unnecessary.

The assignation of "index-figures" for vectors for locomotions other than physical may be accomplished through the operational definition of tension in terms of memory index figures, or in the tendency to resume interrupted acts. (16).

In the *social field,* the relative strength of vectors may be operationally defined through attainment or failure to attain membership-character in groups where the social goal lies within definite social regions or statistically through the outbreak of war, revolution or industrial strike.

V. Hodological Space

Vectors are directed magnitudes and the problem arises as to the definition of direction in psychological fields. Lewin has recently attempted the mathematical solution of this problem, and has been able to show under what conditions direction of vectors may be defined and what the prerequisites to such definition are. The following lines give only his findings. For the mathematical deductions and proofs the reader must go to his original paper.

For physical locomotions where there is no barrier between the initial position and the goal, the problem of definition of direction raises no particular difficulties. Direction is a binary spatial relationship, which may be defined in Euclidean space by two points and their sequence. Hence the direction from point *a* to *b* of a Euclidean plane

is given by the straight line joining them. The direction of the vector underlying physical locomotion, where there is no barrier lying in the line between the organism and the goal is given by the straight line joining the organism and the goal. Hence a child walking towards a piece of candy in a room, is to be ordered to a field as in Figure VIII.

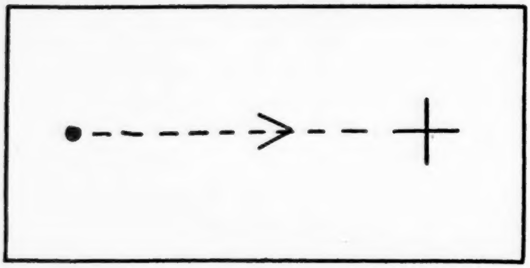


FIGURE VIII

Such situations, however, are of little psychological interest and as soon as a barrier is imposed, the direction of the vector in quasi-physical space is more difficult of definition. Lewin introduces the concept hodological space, (i.e., space of the path) and distinguishes between special and general hodological space. Special hodological space is space in which direction is defined by the initial differential of the distinctive path between two points in the space. By distinctive (ausgezeichnet) path Lewin means one distinguished through some dynamic criterion, such as being the shortest in time, space, energy expenditure, or, under other conditions, longest in time, etc. Consequently, if a barrier is placed between the child and the candy in the above example, the hodological direction is as given in Figure IX. The properties of such a space are immediately dependent on the psychobiological dynamics of the situation, because direction in it is definable only when these factors are taken into consideration. We saw above (Section II), however, that such a procedure is quite allowable in modern geometry. All problems of physical locomotion, where the initial position of the organism and the goal may be ordered to definite points, may be handled in special hodological space. However, in hodological space as opposed to Euclidean, there are multidimensional regions where the points are undifferentiated with regard to direction from a given initial point, and there are point-pairs which are not related by a direction. Thus in Figure X, all the points in the shaded region A lie in the same direction from P_1. (The directions P_{12}, P_{13}, P_{14}, are hodologically identical.) There is further *no* direction be-

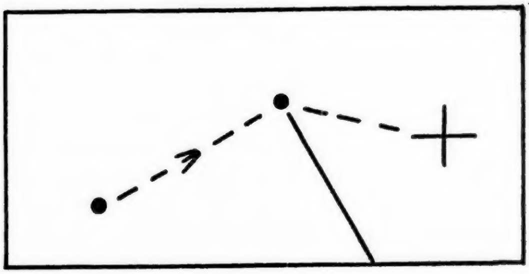
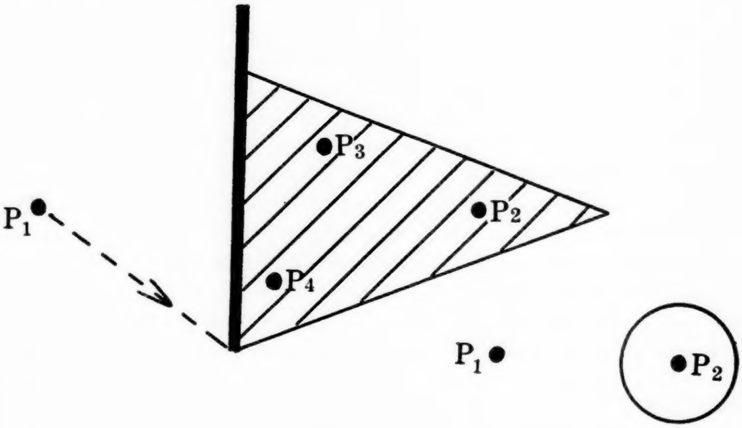


FIGURE IX

tween P_1 and P_2 in Figure XI, as there is no path between them. *The direction in quasi-physical space depends on the properties of the* properties of the total field.



FIGURES X AND XI

When we attempt to define direction for problems where there is no direct correlated physical locomotion (problems in quasi-social and quasi-conceptual space) the difficulties which beset us are even greater. Physical space is usually infinitely structured (durchstruktuiert), while conceptual and social space in general only allow position to be topologically defined, i.e., they are structured, but not infinitely structured.* Consequently the paths in *general* hodological

*Cf. the distinctions structured, unstructured, infinitely structured, given above.

space are between topological regions rather than between points. *Direction in general hodological space is defined as the step from the initial region to that contiguous region, which lies in the distinctive path to the goal.* In the example given above, of the freshman and the fraternity, the direction towards C can be defined through the locomotion A to B. The *direction* in general hodological space is hence *relative* to the degree of structure of the psychological space. Lewin defines psychological space as a general hodological space. Consequently, the magnitude of our vectors may be defined in terms of an index-figure and its *direction* in terms of hodological space.

VI. Existing Applications of the Foregoing Constructs

K. Lewin first applied topological concepts to psychological research. Lewin's consideration (13) of the existing relationship between psychology and the scientific method convinced him of the necessity for an hypothetico-deductive approach. He was further convinced that psychology should use precise, mathematically defined theoretical constructs. For the reasons given above, topological concepts were seen to be the most adequate. Topological concepts alone, however, cannot investigate dynamical processes and Lewin has made wide use of non-metricized dynamical concepts. We also owe to Lewin the development of hodological space.

The chief experimental researches based on this methodology are included in a series of papers entitled "Untersuchungen zur Handlungs-und Affektpsychologie" appearing currently in *Psychologische Forschung.* This series of papers is composed chiefly of dissertations done under Lewin's direction. There is now available in English a translation of certain of Lewin's theoretical papers (16). This last work ends with a chapter which abstracts the individual papers of the *Forschung* series. A larger work of Lewin's is to be published shortly. (14).

Lewin has also suggested the applicability of his method to the problem of sociology and social psychology (14). The writer has attempted to apply this suggestion in a methodological consideration of social psychology (3). At the present time the writer is engaged with a larger work on social psychology from the standpoint of the theory of the social field, which he hopes to publish within the next year.

VII BIBLIOGRAPHY

1. BROWN, J. F., A methodological consideration of the problem of psychometrics. *Erkenntnis,* 1934, *4,* 46-61.
2. BROWN, J. F., Freud and the scientific method. *Phil. of Science,* 1934, *1,* 323-337.
3. BROWN, J. F., Towards a theory of social dynamics. *Jl .of Soc. Psych.,* (in print).
4. BROWN, J. F. and D. D. FEDAR, Thorndike's theory of learning as gestalt psychology. *Psychol. Bull.,* 1934,*31,* 426-437.
5. CAMPBELL, N. R., *Physics: The Elements,* Cambridge: University Press, 1920.
6. CARNAP, R., Die physikalische Sprache als Universalsprache der der Wissenschaft, *Erkenntnis,* 1931, *2,* 432-465.
7. CARNAP, R., Psychologie in physikalischer Sprache. *Erkenntnis,* 1932, *3,* 107-142.
8. FRAENKEL, A. ,*Einleitung in die Mengenlehre.* Berlin: Julius Springer, 1923.
9. FRANKLIN, P. What is topology? *Phil. of Sci.,* 1935, 2, 39-47.
10. HEIDER, F., Ding und Medium. *Symposium,* 1926, *I,* 109 ff.
11. HULL, C., The concept of the habit, family hierarchy and maze learning, *Psychol. Rev.,* 1934, *42,* 35-52, 134-152.
12. KERÉKJÁRTÓ, B., Vorlesungen über Topologie. I, *Flachentopologie.* Berlin: Julius Springer, 1923.
13. LEWIN, K. The conflict between Aristotelian and Galilean modes of thought in psychology. *Jl. Gen. Psych.,* 1931, *5,* 141-177.
14. LEWIN, K., *Die Grundlagen der dynamischen Psychologie.* (To be published, 1935.)
15. LEWIN, K., Der Richtungsbegriff in der Psychologie. *Psychol. Forsch.,* 1934, *19,* 250-299.
16. LEWIN, K., A dynamic theory of personality. Translated by D. Adams and K. Zener. New York: McGraw-Hill, 1935.
17. POINCARÉ, H., Analysis Situs, *Journal de l'Ec. Pol.,* 1895, Vol. II, Part I, 1-121.
18. RIEMANN, B. *Über die Hypothesen, welche der Geometrie zu Grunde liegen.* Edited by H. Weyl. Berlin: Julius Springer, 1923.
19. THORNDIKE, E. L., *The measurement of intelligence.* New York, 1927.
20. TOLMAN, E. C., and BRUNSWIK, E., The organism and the causal texture of the environment. *Psychol. Rev.,* 1935, *42,* 43-77.

PSYCHOMETRIKA—VOL. 1, NO. 2
JUNE, 1936

RELIABILITY COEFFICIENTS IN A CORRELATION MATRIX

SAMUEL A. STOUFFER

University of Chicago

Given s fallible tests $t_1, t_2, \cdots t_s$, the problem is to express their intercorrelations in terms of the average correlations between a varying number of parallel forms contained within each test. A new correlation determinant Δ' is derived containing d_{ii} instead of unity as an element on the principal diagonal, where

$$d_{ii} = [1 + (m_i - 1)\overline{r}_{ii}]/m_i ,$$

in which m_i is the number of parallel forms comprising the i'th test and $\overline{r}_{ii}$ is the average intercorrelation of the $m_i(m_i - 1)/2$ parallel forms. As $m_i \to \infty$, d_{ii} approaches the correlation "corrected for attenuation." These results make explicit the assumptions, as to intrinsic accuracy of all measures, which are implicit in the usual multiple and partial correlation analysis. These results also make possible a simple procedure for estimating the effect on various partial correlation measures of improving the accuracy of part or all of the measures by including additional parallel forms.

Given s fallible tests $t_1\ t_2, \cdots, t_s$, to express the conventional correlation determinant

$$\Delta = \begin{vmatrix} 1 & r_{t_1t_2} & \cdots & r_{t_1t_s} \\ r_{t_1t_2} & 1 & \cdots & r_{t_2t_s} \\ \cdot & \cdot & \cdot & \cdot \\ r_{t_1t_s} & r_{t_2t_s} & \cdots & 1 \end{vmatrix} \tag{1}$$

in terms of the intercorrelations between parallel forms comprising each test t_i. Write

$$\begin{aligned} t_1 &= z_{1_1} + z_{1_2} \cdots + z_{1_k} \cdots + z_{1_{m_1}} \\ t_2 &= z_{2_1} + z_{2_2} \cdots + z_{2_k} \cdots + z_{2_{m_2}} \\ &\cdots\cdots\cdots\cdots \\ t_i &= z_{i_1} + z_{i_2} \cdots + z_{i_k} \cdots + z_{i_{m_i}} \\ &\cdots\cdots\cdots\cdots \\ t_s &= z_{s_1} + z_{s_2} \cdots + z_{s_k} \cdots + z_{s_{m_s}} , \end{aligned}$$

where each z_{i_k} is one of m_i parallel forms comprising test t_i and is expressed in standard measure, that is,

$$z_{i_k} = (X_{i_k} - \overline{X}_{i_k})/\sigma_{i_k} ,$$

in which X_{i_k} is a raw score and $\overline{X}_{i_k}$ and σ_{i_k} are the mean and standard

deviation, respectively, of the raw scores. Since

$$\sigma^2 z_{i_k} = 1\,,\ \sigma^2 t_i = m_i + 2(r_{i_1 i_2} + r_{i_1 i_3} \cdots + r_{i_{m_i-1} i_{m_i}})$$
$$= m_i[1 + (m_i - 1)\bar{r}_{ii}]\ ,$$

where

$$\bar{r}_{ii} = 2(r_{i_1 i_2} + r_{i_1 i_3} \cdots + r_{i_{m_i-1} i_{m_i}})/m_i(m_i - 1)\ ,$$

the average of the $m(m-1)/2$ reliability coefficients.

Write $r_{t_i t_j} = \sum t_i\, t_j/n\, \sigma_{t_i}\, \sigma_{t_j}$, where n is the number of individuals taking the tests. Since

$$\sum z_{i_k} z_{j_k}/n = r_{i_k j_k}\ ,\ \sum t_i\, t_j/n = r_{i_1 j_1} + r_{i_1 j_2} \cdots + r_{i_1 j_{m_j}} \cdots$$
$$+ r_{i_{m_i} j_1} + r_{i_{m_i} j_2} \cdots + r_{i_{m_i} j_{m_j}} = m_i\, m_j\, \bar{r}_{ij}\ ,$$

where $\bar{r}_{ij}$ is the average of the $m_i\, m_j$ intercorrelations. We then have,

$$r_{t_i t_j} = \frac{m_i\, m_j\, \bar{r}_{ij}}{\sqrt{m_i m_j[1 + (m_i - 1)\bar{r}_{ii}][1 + (m_j - 1)\bar{r}_{jj}]}}$$

$$= \frac{\bar{r}_{ij}}{\sqrt{\left[\dfrac{1 + (m_i - 1)\bar{r}_{ii}}{m_i}\right]\left[\dfrac{1 + (m_j - 1)\bar{r}_{jj}}{m_j}\right]}}$$

$$= \frac{\bar{r}_{ij}}{\sqrt{d_{ii}\, d_{jj}}}\ , \tag{2}$$

where $d_{ii} = [1 + (m_i - 1)\bar{r}_{ii}]/m_i$.*

*Equation 2 is identical, though it is expressed in different notation, with (147) in Truman L. Kelley, *Statistical Method*, page 197. It is assumed that each parallel form comprised in t_i has unit weight. If the m_i forms are assigned varying weights w_k, $(k = 1, 2, \cdots m_i)$, Kelley's (149), page 198, may be used.

Substituting (2) in (1) we have

$$\Delta = \begin{vmatrix} 1 & \dfrac{\bar{r}_{12}}{\sqrt{d_{11}\, d_{22}}} & \cdots & \dfrac{\bar{r}_{1s}}{\sqrt{d_{11}\, d_{ss}}} \\ \dfrac{\bar{r}_{12}}{\sqrt{d_{11}\, d_{22}}} & 1 & \cdots & \dfrac{\bar{r}_{2s}}{\sqrt{d_{22}\, d_{ss}}} \\ \cdot & \cdot & \cdot & \cdot \\ \dfrac{\bar{r}_{1s}}{\sqrt{d_{11}\, d_{ss}}} & \dfrac{\bar{r}_{2s}}{\sqrt{d_{22}\, d_{ss}}} & \cdots & 1 \end{vmatrix} \tag{3}$$

$= \Delta'/d_{11}\, d_{22} \cdots d_{ss}$, where (4)

$$\Delta' = \begin{vmatrix} d_{11}\, \bar{r}_{12} & \cdots & \bar{r}_{1s} \\ \bar{r}_{12}\, d_{22} & \cdots & \bar{r}_{2s} \\ \cdot\ \cdot & \cdot & \cdot \\ \bar{r}_{1s}\, \bar{r}_{2s} & \cdots & d_{ss} \end{vmatrix} ,$$

while any $(s - 1)$-rowed minor of Δ ,

$$\Delta_{ij} = \frac{\sqrt{d_{ii}\, d_{jj}}}{d_{11}\, d_{22} \cdots d_{ss}} \Delta'_{ij} ,$$

where Δ'_{ij} is the corresponding minor of Δ'.

As all values $r_{ii} \to 1$; and as all values $\bar{r}_{ij} \to r_{i_k j_k}$, the correlation between any two parallel forms in t_i and t_j; Equations 3 and 4 approach the form

$$\Delta = \begin{vmatrix} 1\ r_{12} & \cdots & r_{1s} \\ r_{12}\ 1 & \cdots & r_{2s} \\ \cdot\ \cdot & \cdot & \cdot \\ r_{1s}\, r_{2s} & \cdots & 1 \end{vmatrix} .$$

As m_i, the number of parallel forms of the i-th test, $\to \infty$ (so that $1/m \to 0$),

$$d_{ii} = \left[\frac{1}{m_i} + \left(1 - \frac{1}{m_i}\right) \bar{r}_{ii} \right] \to \bar{r}_{ii} \text{ , whence from (2),}$$

$$r_{t_i t_j} \to \frac{\bar{r}_{ij}}{\sqrt{\bar{r}_{ii}\, \bar{r}_{jj}}} ,$$

a form of the correlation between measures of the i'th and j'th tests "corrected for attenuation," permitting us to write, from (4)

$$\Delta = \Delta''/\bar{r}_{11}\, \bar{r}_{22} \cdots \bar{r}_{ss} , \quad (5)$$

where

$$\Delta'' = \begin{vmatrix} \bar{r}_{11}\, \bar{r}_{12} & \cdots & \bar{r}_{1s} \\ \bar{r}_{12}\, \bar{r}_{22} & \cdots & \bar{r}_{2s} \\ \cdot\ \cdot & \cdot & \cdot \\ \bar{r}_{1s}\, \bar{r}_{2s} & \cdots & \bar{r}_{ss} \end{vmatrix} .$$

It is thought that these relationships may help to make explicit some of the assumptions implicit in the use of test measures in correlation analysis, as well as to provide a practical technique for estimating the probable effects on final correlation results of improving a part or all of the tests by including additional parallel forms.*

Examples of Derived Measures when $s = 3$. From Equation 4, when $s = 3$, we have

$$r_{t_1 t_2 \cdot t_3} = \frac{\Delta_{12}}{\sqrt{\Delta_{22}\,\Delta_{11}}} = \frac{\Delta'_{12}}{\sqrt{\Delta'_{22}\,\Delta'_{11}}} = \frac{\bar{r}_{12}\,d_{33} - \bar{r}_{13}\,\bar{r}_{23}}{\sqrt{(d_{11}\,d_{33} - \bar{r}^2_{13})\,(d_{22}\,d_{33} - \bar{r}^2_{23})}} \tag{6}$$

$$\beta_{t_1 t_2 \cdot t_3} = \frac{\Delta_{12}}{\Delta_{11}} = \frac{\Delta'_{12}}{\Delta'_{11}}\sqrt{\frac{d_{22}}{d_{11}}} = \left[\frac{\bar{r}_{12}\,d_{33} - \bar{r}_{13}\,\bar{r}_{23}}{d_{22}\,d_{33} - \bar{r}^2_{23}}\right]\sqrt{\frac{d_{22}}{d_{11}}} \tag{7}$$

$$R^2_{t_1 \cdot t_2 t_3} = 1 - \frac{\Delta}{\Delta_{11}} = 1 - \frac{\Delta'}{\Delta'_{11}\,d_{11}} = \frac{\bar{r}^2_{12}\,d_{33} + \bar{r}^2_{13}\,d_{22} - 2\,\bar{r}_{12}\,\bar{r}_{13}\,\bar{r}_{23}}{d_{11}(d_{22}\,d_{23} - \bar{r}^2_{23})} \tag{8}$$

Equations such as these may prove valuable not only in cases where d_{ii} and $\bar{r}_{ij}$ are known, but also in cases where it is desired to insert various guessed values of d_{ii} and $\bar{r}_{ij}$. As m_1, m_2 and m_3, the number of parallel forms, approach ∞, $d_{ii} \rightarrow \bar{r}_{ii}$, from (5), permitting us to rewrite (6), for example, as

$$r_{t_1 t_2 \cdot t_3} = \frac{\bar{r}_{12}\,\bar{r}_{33} - \bar{r}_{13}\,\bar{r}_{23}}{\sqrt{(\bar{r}_{11}\,\bar{r}_{33} - \bar{r}^2_{13})\,(\bar{r}_{22}\,\bar{r}_{33} - \bar{r}^2_{23})}}\,, \tag{9}$$

a useful form provided that we feel reasonably safe in our estimates of $\bar{r}_{ii}$ and $\bar{r}_{ij}$.

*This paper presents a further generalization of results obtained by setting each $t_i = z_i + z'_i$, as reported by the writer in "Evaluating the Effect of Inadequately Measured Variables in Partial Correlation Analysis," *Journal of the American Statistical Association*, June, 1936. Applications given in this paper make use of sociological and economic data, though it would be very easy to find examples in the psychological field.

PSYCHOMETRIKA—VOL. 1, NO. 2
JUNE, 1936

FURTHER CONTRIBUTIONS TO THE MATHEMATICAL THEORY OF HUMAN RELATIONS

N. RASHEVSKY

University of Chicago, Chicago, Illinois

In continuation of a previous paper, some consequences of the fundamental equations established there are studied. For some simple hypothetical cases it is shown how some of the parameters which enter in the equations governing the structure of the social group can be determined by means of those equations from actually observable data. Furthermore some general properties of the variation with respect to time of the fundamental distribution function, which enters in the equations, are derived.

I.

In a previous paper* we have outlined a general mathematical approach to a deductive, theoretical science of human society. We have established various general forms of equations which can be used for that purpose, and indicated the way of describing in mathematical terms some phenomena of social interactions. In those fundamental equations we have been making use of some quantities which directly are not accessible to physical measurements. We justified however this procedure by invoking the example of physical sciences, where equations are established between quantities which may directly be inaccessible to measurements, but which yet are measurable "in principle". This "measurability in principle" is made possible by the application of the mathematical equations, which eventually give us connections between those directly unmeasurable quantities and others which can be measured directly. It is the purpose of the present paper to extend the preceding study and to illustrate by means of a few examples the way by which in the ultimate development of the theoretical system, a measurement and determination "in principle" of the quantities involved can be achieved. We shall in other words, derive some special consequences from some of the general equations, and show how those can be in principle compared to observable data.

We must however very strongly emphasize again, that such a comparison in principle does not mean a comparison with *actual existing data*. We are here concerned with a purely deductive, theoretical science, and we therefore study at first purely imaginary cases,

*Phil. of Sc., 1935, 2, 413, hereinafter referred to as *loc. cit.*

which due to the intentional oversimplification, have no real existence. If we speak of a comparison of an equation with *observable* data, we meant thereby just this: we consider a simple, *theoretically possible*, but not an actual case, and set up equations which describe it. The equations themselves may be of such a nature, that even if this case studied would really exist, they would not be directly verifiable. But some of their consequences could be compared with observable data, *which would be available if our case really existed.*

Inasmuch as even the simplest actual cases are much more complex than the imaginary cases with which we perforce must begin, no claim whatsoever can be made *at this stage* for any practical application to actual cases. In the long run however after a systematic study of more and more complex cases, we may arrive at a representation of some actual ones. But this will of necessity require a great deal of preliminary, purely abstract theoretical study. This has been the way of all exact science.*

II.

We shall begin by considering the simplest case of one activity only, represented by equ. 5, *loc cit.*, and first simplify the case still further by considering that all individuals have the same desire w for the activity a_i, and differ only in their coefficients of influence F. The activity a_i itself may be of any kind: production of food or other practical necessities, or any artistic activity, like painting, composing, etc. The proper practical unit for a_i may vary from case to case. But in principle for a given activity its intensity a_i may be expressed by the amount of energy spent on it, on the average, per unit of time. Since in the absence of the influence of others we have:

$$a_i = a\,w \ , \tag{1}$$

where a is the coefficient of proportionality, dependent on the choice of units, we may measure w by the amount of a_i, which an individual does of his own free will. Then the dimensions of both a_i and w are

$$[m]\,[l]^2[t]^{-3} \tag{2}$$

and, since β is also only a coefficient of proportionality, F becomes a pure number. Putting in equation (5) of *loc. cit.* $a = \beta = 1$ and remembering that in the present case $w = w'$, we find for two individuals

$$a_i = w[1 + (F' - \mathrm{F})] \ . \tag{3}$$

*Cf. our paper in Phil. of Sc., **1**, 176, 409, 1934, **2**, 73, 1935. Nature, **135**, 528, 1935. Biol. Rev. **11**, 345, 1936.

We see that the difference of the coefficient of infiuence of two individuals equals to 1, if one of them can through his influence on another induce him to double the amount of the latter's activity done of free will.

Of course we must remember that this holds only for the particular hypothetical case which we decided to consider here. Adoptions of different postulates will require a modification of the set-up of units, but the procedure remains fundamentally the same. In subsequent publications we intend to make a systematic investigation of all possible postulates discussed in loc. cit., as well as of some different ones.

III.

The next question is the choice of a function $N(F, w)$, which in our case degenerates into $N(F)$. Again the proper procedure would be to investigate various possible $N(F)$, and to begin with the simplest cases. One may think that a normal distribution should be investigated first. In view of its wide range of applications in statistics, a normal distribution certainly deserves a separate study. In view however of the fact that the evaluation of some integrals in finite form becomes impossible for a normal distribution, we shall consider here a different one, namely:

$$N(F) = AFe^{-aF} \ , \tag{4}$$

A and a being positive constants. This $N(F)$ is equal to zero for $F = 0$ and $F = \infty$, has a maximum for $F = 1/a$, and an inflection point for $F = 2/a$.

From the requirement that

$$\int_0^\infty N(F)\,dF = \mathfrak{N} \ , \tag{5}$$

where $\mathfrak{N}$ is the total number of individuals, we find

$$A = \mathfrak{N}\, a^2 \ , \tag{6}$$

so that

$$N(F) = \mathfrak{N} a^2 F\, e^{-aF} \ . \tag{7}$$

It may be argued that the integration in (5) should be carried out not from zero to infinity, but from zero to F_m, F_m being the maximum value of F that occurs in the population. This would give

$$A = \mathfrak{N}/[1/a^2 - (F_m + 1/a)\,e^{-aF_m}/a] \ , \tag{8}$$

which reduces to (6), if F_m is so large that e^{-aF_m} is very small. In

this paper we shall confine ourselves to such a case, and use (6) instead of (8), though the whole theory may be developed in quite a similar way on the basis of (8), leading only to somewhat more complex formulae.

Of particular interest is the case where only one individual in the group has the highest F_m. We then have

$$N(F_m) = \mathfrak{N} a^2 F e^{-aF_m} = 1 \ , \tag{9}$$

or

$$\log \mathfrak{N} + \log a + \log aF_m - aF_m = 0 \ . \tag{10}$$

Neglecting log aF_m as compared with aF_m, we find approximately

$$F_m = (\log \mathfrak{N} + \log a)/a \ . \tag{11}$$

IV.

With the shape of $N(F)$ chosen, let us turn to the equation governing the formation of "social classes" discussed in section IV of *loc. cit.* Taking as the criterion of "association" equation 16 of *loc. cit.*,

$$(F' - F)^2 < \Delta^2 \ , \tag{12}$$

the lower limit F_x for the "upper class" is given by the root of the equation

$$\int_{F_x}^{F_m} \int_{F_x}^{F_m} [(F' - F)^2 - \Delta^2] \, N(F) N(F') dF dF' = 0 \ . \tag{13}$$

Introducing into (1) for $N(F)$ and $n(F')$ the expression (7) we find after somewhat elaborate calculations:

$$\begin{aligned}
&\mathfrak{N}^2 a^2 \{ e^{-2aF_x} [(\frac{2}{a^2} - \Delta^2) F_x^2 + (\frac{8}{a^2} - \frac{2\Delta^2}{a}) F_x + \frac{4}{a^4} - \frac{\Delta^2}{a^2}] \\
&\quad + e^{-2aF_m} [(\frac{2}{a^2} - \Delta^2) F_m^2 + (\frac{8}{a^3} - \frac{2\Delta^2}{a}) F_m + \frac{4}{a^4} - \frac{\Delta^2}{a^2}] \\
&\quad - 2 e^{-a(F_x+F_m)} [(F_m + \frac{1}{a}) F_x^3 - (\frac{F_m}{a} + \frac{1}{a^2} + 2 F_m^2) F_x^2 \\
&\quad + (F_m^3 - \frac{2F_m^2}{a} + \frac{4F_m}{a^2} - \Delta^2 F_m + \frac{4}{a^3} - \frac{\Delta^2}{a}) F_x + \frac{4F_m}{a^3} \\
&\quad - \frac{\Delta^2 F_m}{a} + \frac{4}{a^4} - \frac{\Delta^2}{a^2} + \frac{F_m^3}{a} - \frac{F_m^2}{a^2}]\} = 0 \ .
\end{aligned} \tag{14}$$

An exact solution of this transcendental equation is rather difficult, but we have an approximate expression for the case that Δ is sufficiently small and therefore F_x is large, however so that $F_x << F_m$. Then we may neglect the term multiplied by e^{-2aF_m} as compared with those multiplied by e^{-2aF_x}, and also keep in the braces only terms in F_x^3 and F_x^2. This gives

$$\left(\frac{2}{a^2} - \Delta^2\right) F_x^2 \, e^{-2aF_x} - 2\left(F_m + \frac{1}{a}\right) F_x^3 \, e^{-a(F_x+F_m)} = 0 \ , \tag{15}$$

or after transformation,

$$\log\left(\frac{2}{a^2} - \Delta^2\right) - \log 2\left(\frac{F_m}{a} + \frac{1}{a^2}\right) - aF_x$$

$$- \log aF_x + aF_m = 0 \ , \tag{16}$$

and again neglecting $\log aF_x$ as compared with aF_x, we finally obtain:

$$F_x = F_m - a \log \frac{2(F_m/a + 1/a^2)}{2/a^2 - \Delta^2} \ . \tag{17}$$

Since $2(F_m/a + 1/a^2) = 2/a^2 + 2F_m/a > 2/a^2 - \Delta^2$, the log in (17) is positive and $F_x < F_m$. With increasing Δ^2, F_x decreases, as should be the case.

V.

Another relation involving F_x is obtained by considering the ratio ϑ of the number of individuals of the upper class, to the total number of individuals in the society. This is given by

$$\vartheta = \int_{F_x}^{\infty} N(F)\, df \Big/ \int_0^{\infty} N(F)\, dF \ . \tag{18}$$

Introducing (7) into (18) we find:

$$\vartheta = a\left(F_x + \frac{1}{a}\right) e^{-aF_x} \ , \tag{19}$$

or

$$\vartheta = (aF_x + 1)\, e^{-aF_x} \ . \tag{20}$$

Equ. (20) gives a relation between F_x, a and ϑ which latter is a directly measurable quantity. For the case $aF_x > 1$, (20) may be simplified thus:

$$\vartheta = aF_x e^{-aF_x} \ ,$$

or

$$\log \vartheta = \log aF_x - aF_x \ .$$

Neglecting again log aF_x as compared with aF_x we find

$$F_x = -\frac{1}{a} \log \vartheta \,. \tag{21}$$

F_x is always positive, because $\vartheta < 1$.

A still further relation is provided between F_x, a and w, which we consider as constant, by calculating the total amount A of the activity of the whole group:

$$A = \int_0^\infty a(F) N(F) dF = w \int_0^\infty N(F) dF + w \int_0^\infty N(F) dF$$

$$\times \int_F^\infty N(F')(F' - F) dF' \,, \tag{22}$$

which can also be easily calculated by using (17). The constant desire w of every individual being also measurable in principle, we have 4 equations: (11), (17), (20) and (22) by means of which we can express F_m, F_x, a and Δ in terms of ϑ, A and w. We see thus how the various quantities, which we introduced into our fundamental equations and which directly cannot be measured, can nevertheless be determined by means of the equations themselves, from other directly measurable quantities. Similar investigations will be carried out for more complex cases, involving more variables.

VI.

Now let us investigate somewhat closer the function $N(F)$ itself. In *loc. cit.* we have considered the variation of $N(F)$ due to the fact that the progeny of an individual with a definite F may have in general a different value of F. Let again $p(F^*, F)$ be the number of individuals having a characteristic F and born of parents F^*. In general we must consider the case, that the two parents have a different F, but for the time being we confine ourselves to the simpler case, that both parents have an identical F. The total number of individuals with characteristic F, born of any parents per unit time is (*loc. cit.*, equation 29)

$$\int_0^\infty n(F^*) N(F^*, t) p(F^*, F) dF^* \,, \tag{23}$$

where $n(F^*)$ denotes the birth rate per individual. If $m(F)$ is the death rate also per individual, we have for the total change of $N(F,t)$ per unit time

$$\frac{\partial N(F,t)}{\partial t} = \int_0^\infty n(F^*) N(F^*,t) p(F^*,F) dF^* - m(F) N(F) \; , \tag{24}$$

Let us consider the simplest case, that both $n(F)$ and $m(F)$ are constants, that is that the birth and death rates are the same for all types of individuals. Then (24) becomes:

$$\frac{\partial N(F,t)}{\partial t} = n \int_0^\infty N(F^*,t) p(F^*,F) dF^* - mN(F) \; . \tag{25}$$

We shall solve equation (25) by putting

$$N(F,t) = N^*(F) \varphi(t) \; , \tag{26}$$

where $N^*(F)$ is a function of F only, and $\varphi(t)$ is a function of t only, and determining $N_0(F)$ and $\varphi(t)$ so as to satisfy both equation (25) and the requirement that an initial moment t, which we may put without any loss of generality equal to zero, $N(F,t)$ should be a given function $N_0(F)$ of F.

Introducing (26) into (25) we find:

$$N^*(F) \frac{d\varphi}{dt} = n \varphi(t) \int_0^\infty N^*(F^*) p(F^*,F) dF^* - mN^*(F) \varphi(t) \; , \tag{27}$$

or putting

$$\frac{n \int_0^\infty N^*(F) p(F^*,F) dF^*}{N^*(F)} - m = a \; , \tag{28}$$

$$\frac{d\varphi}{dt} - = a \varphi \; , \tag{29}$$

which gives

$$\varphi = A e^{at} \; , \tag{30}$$

A being a constant of integration.

For every given value of a, equation (28) gives us an equation for the determination of $N^*(F)$. But (28) can be written after simple rearrangements in the following way:

$$N^*(F) = \frac{n}{a+m} \int_0^\infty N^*(F^*) p(F^*,F) dF) \; , \tag{31}$$

which is a homogeneous integral equation of second kind, with the kernal $p(F^*,F)$. Equ. (31) possesses solutions only for definite values of the constant

$$\lambda = \frac{n}{a+m} \cdot \tag{32}$$

Let $\lambda_1, \lambda_2, \cdots, \lambda_i$ be those "eigenvalues" arranged in increasing order so that:

$$\lambda_1 < \lambda_2 < \lambda_3 < \cdots . \tag{33}$$

Then in order that (31) should have solutions at all, a must have one of the values

$$a_i = \frac{n - \lambda_i m}{\lambda_i} = \frac{n}{\lambda_i} - m , \tag{34}$$

If $N_i^*(F)$ is an eigenfunction of (31), corresponding to the eigenvalue λ_i then:

$$A_1 N_i^*(F) e^{a_i t}$$

is a particular solution of (25), and the general solution is given by:

$$N(F,t) = \sum_1^\infty A_i N_i^*(F) e^{a_i t} . \tag{35}$$

For $t = 0$, this is equal to

$$N(F,0) = \sum_1^\infty A_i N_i^*(F) , \tag{36}$$

and since all $N_i^*(F)$ form a complete orthogonal system, the coefficients A_i can be determined so, that

$$\sum_1^\infty A_i N_i^*(F) = N_0(F) . \tag{37}$$

We have

$$A_i = \int_0^\infty N_0(F) N_i^*(F) dF . \tag{38}$$

Equations (35) and (38) represent the general solution of (25). Let us consider some of the consequences.

On account of (33) we have:

$$a_1 > a_2 > a_3 > \cdots > \; ; \; a_\infty = -m . \tag{39}$$

If

$$m > \frac{n}{\lambda_1} \tag{40}$$

then all a's are negative. In that case, regardless of the choice of the coefficients A_i, in other words regardless of the initial distribution

$N_0(F)$, the expression $N(F,t)$ given by equation (35) tends to zero. Expression (40) sets, therefore an upper limit for the death-rate m, above which the social group will with time become extinct.

If however

$$m < \frac{n}{\lambda_1} \tag{40a}$$

then some α's, say $\alpha_1, \alpha_2 \cdots \alpha_s$ will be positive, others $\alpha_{i+1}, \alpha_{s+2} \cdots$ will be negative. But then all terms in (35) above the s-th will tend to zero, while the first s terms will increase. But with increasing t the term with the largest α, that is the first term, will exceed the others more and more, the ratio

$$e^{\alpha_i t}/e^{\alpha_1 t} \qquad (i \leqq s)$$

tending to zero. Hence after a sufficient time has elapsed, $N(F,t)$ will be given by

$$N(F,t) = A_1 N_1^*(F) e^{\alpha_1 t} \ . \tag{41}$$

Equation (41) shows, that the total number $\mathfrak{N}$ of individuals will increase exponentially, but the distribution $N(F)$ will not vary, being given by the first eigenfunction $N_1^*(F)$ of the integral equation (31). A similar result holds for the special case that $m = \dfrac{n}{\lambda_1}$. Then $\alpha_1 = 0$, all others are negative, and (35) tends asymptotically to $A_1 N_1^*(F)$. In this case not only does the distribution function tend asymptotically to a fixed form, but the total number of individuals also tends to be constant.

We thus find a fundamental result: under the simplified assumptions made here, the distribution function $N(F)$ tends always either to zero or to a stationary distribution, which is determined by the function $p(F^*,F)$, since the latter is the kernel of the integral equation (31), whose first eigenfunction determines the stationary distribution. Any disturbances, like wars, starvations, etc., may upset this distribution temporarily, but in time it will again be restored. In a subsequent paper we shall investigate a special case of $p(F^*, F)$ and the resulting $N(F)$.

In *loc. cit.* we have seen, how the variation of $N(F)$ with time determines the variation of the social structure and causes eventually its instability. Observations of the variations in time of the social structure may also lead us to equations which connect some of the directly unobservable quantities with directly measurable ones. For instance, as we have seen in *loc. cit.*, due to variation of $N(F)$, the

second class will contain a certain number of individuals with $F > F_z$. If the variation of $N(F)$ with respect to time is given, then this number R of individuals with $F > F_z$ is also given for any moment, $R = R(t)$. But these individuals will not be controlled in a normal way by the first class. If, as is usually the case, such controlling class uses various methods of coercion against active political opponents, then $R(t)$ would represent the number of individuals subject to such coercion, and this number can be directly determined. If this number is N_c, then

$$R(t) = N_c \tag{42}$$

gives us an equation, involving some parameters, which determine the variation of $N(F)$, $p(F^*,F)$, etc. Together with other possible equations it may be used to calculate those parameters.

VII.

In the general case the coefficient of influence may itself be a derivative notion. If an individual can easily perform and does perform an activity the results of which are badly needed by another individual then the first individual may have a strong control over the second. We may thus consider the coefficients of influence as being functions of the activities. For the case of two activities the situation is mathematically represented in the following manner.

Let each individual be characterized by the desires w_1 and w_2 for the performance of the activities a_1 and a_2, and by the desires u and u_2 to possess the results of the corresponding activities, without actually performing them:

Then:

$$\left.\begin{aligned} &a_1(w_1,u_1,w_2,u_2) = aw_1 \\ &+ \beta u_2 \int_0^\infty N(w'_1,u'_1,w'_2,u'_2)\, a_2(w'_1,u'_1,w'_2,u'_2)\, u'_1 dw'_1 du'_1 dw'_2 du'_2 \\ &a_2(w_1,u_1,w_2,u_2) = aw_2 \\ &+ \beta u_1 \int_0^\infty N(w'_1,u'_1,w'_2,u'_2)\, a_1(w'_1,u'_1,w'_2,u'_2)\, u'_2 dw'_1 du'_1 dw'_2 du'_2 \end{aligned}\right\} \tag{43}$$

The integrals in the right-hand side of both equations are constants, which we denote by A_1 and A_2 respectively.

$$\left.\begin{aligned} a_1 &= aw_1 + A_1 \beta u_2 \ . \\ a_2 &= aw_2 + A_2 \beta u_1 \ . \end{aligned}\right\} \tag{44}$$

The constants A_1 and A_2 are determined from the two equations

$$\left.\begin{aligned} A_1 &= \int_0^{\infty} N(w_1,u_1,w_2,u_2)\, u_1 (aw_2 + A_2 \beta u_1)\, dw_1 du_1 dw_2 du_2 \ , \\ A_2 &= \int_0^{\infty} N(w_1,u_1,w_2,u_2)\, u_2 (aw_1 + A_2 B u_2)\, dw_1 du_1 dw_2 du_2 \ . \end{aligned}\right\} \quad (45)$$

PSYCHOMETRIKA—VOL. 1, NO. 2
JUNE, 1936

THE RELATION BETWEEN THE DIFFICULTY AND THE DIFFERENTIAL VALIDITY OF A TEST

M. W. RICHARDSON
University of Chicago, Chicago, Illinois

Using scores of 1200 students on a long test as a criterion, each of five subtests of different difficulty has maximum correlation with the criterion when the criterion is dichotomized at a value appropriate to the difficulty of the subtest. A 50-item test element is scored on an all-or-none basis with different standards for passing, and the percentage of passes for successive points on the criterion variable is computed. The Constant Method is applied to this relationship. The limen thus computed is a measure of difficulty, the dispersion is a measure of average (or total) validity, and the slope of the curve is a measure of differential validity. The difficulty of a test element is thus directly related to the maximum differential validity.

I. EXPERIMENTAL DATA

The problem set in this investigation is the relationship between the difficulty of a test or test element and the degree of ability best discriminated by the test. The emphasis is properly placed upon the accurate description of difficulty and validity by the use of a simple mathematical function which will describe both in the same setting. Moreover, this function is to be brought into significant relationship with the statistical methods commonly employed in the interpretation of test results.

The test used in this investigation was the Co-operative General Culture Test, Form 1933.* The test is composed of 803 items, and requires 180 minutes. The choice of this test was based on the following requirements: (1) complete objectivity, (2) freedom from faulty items, (3) length, (4) possibility of getting a large sample from a representative larger population, (5) a wide distribution of difficulty of items. These requirements were approximately met by the Cooperative Test. The subjects were 1,200 college sophomores, a random sample from the 8,996 included in the norms provided by the Committee on Educational Testing. The data, on punched cards, were used at this time to eliminate faulty or completely invalid items. Using total score on the test as a criterion, the population was divided into twelve criterion groups of one hundred each. For each item, the percentage of correct response was computed for each of the twelve

*The 1933 College Sophomore Testing Program. *Report by the Committee on Educational Testing*. Washington: American Council on Education, 1933.

groups. These twelve percentages were plotted against the corresponding average criterion scores of the twelve groups. The slope of this empirical curve was taken as a crude measure of item validity. Items with flat or negatively sloping curves were not used in the experimental subtests. This impressionistic inspection of the validity of the items resulted in the elimination from the experimental tests of thirty or more items of doubtful characteristics. These items were allowed to remain in the criterion, their effect being assumed as negligible because of the length of the total test. From the 780 items remaining, five subtests were selected. The composition of these subtests is described in the following table:

TABLE I

COMPOSITION OF EXPERIMENTAL SUBTESTS

Subtest	Description	Percentage of Correct Answers	Number of Items
A	Very easy	78-95	50
B	Easy	60-77	50
C	Average	41-59	50
D	Difficult	23-40	50
E	Very difficult	5-22	50

The choice of items for the subtests was based on the following considerations, in order of importance: (1) absence from defect, (2) percentage of correct response, (3) balanced positions in the original test, so as not to give disproportionate emphasis to any type of material. To each of the 1,200 subjects a score was given upon each of the five subtests. This scoring was done by machine methods to insure a high order of accuracy.

The scores on the total test were taken as the "criterion" for purposes of this study. The raw scores were converted to normalized standard scores, by groups of 50 subjects. The criterion score of the fifty subjects in each of the twenty-four groups was considered to be the centroid as computed by the formula

$$\bar{x} = \frac{z_1 - z_2}{p_2 - p_1},$$

where $\bar{x}$ is the centroid of a truncated segment of the normal curve with unit standard deviation, z_1 and z_2 are the ordinates enclosing the segment at the left and right respectively, and p_1 and p_2 are the proportions of the area under the curve from the left to the ordinates

with the same subscripts. This procedure is merely a device for dividing the normalized criterion scores into a usable number of class-intervals.

II. Predicting to a Criterion of two Categories

For many purposes we are interested chiefly in estimating from the test scores to which of two categories of some predicted variable the subject belongs. For instance, the testee is "qualified" or "not qualified" on an employment test; he is "passed" or "failed" on a test of achievement. The variance of criterion measures within each of the two categories is thus often neglected in practice. If the score meets a standard specified for the purpose of the test, it may not matter how much individuals, all meeting the standard, differ among themselves. The characteristic of the test which affords discrimination between individuals at different parts of the criterion scale will be referred to as "differential validity" in the subsequent discussion. It is pertinent to inquire the extent to which tests pitched at different levels of difficulty differ in the property of allocating the subject to two categories.

It might be expected that a sufficiently long test, homogeneous as to difficulty of test elements, would approach a condition of uniform differential validity for the total range of the criterion. This argument is based on probability theory, and there is not available any empirical evidence which could be used to test such a generalization. For a uniform difficulty of items which is not $p = .5$, the distribution of scores will be skew, for small and moderate values of n, the number of items. It would be possible, with much labor, to test empirically the degree of skewness of the distribution of scores of very long tests of homogeneous difficulty other than 50 per cent. It is not probable that the skewness diminishes to negligible values for tests of any practicable length. Associated with this skewness, there should be differences in the validity of test for predicting into two categories. It is possible to determine the degree to which the five tests of equal length but of varying difficulty are valid for predicting into two categories. The criterion was systematically divided into two categories at twenty-three different points indicated in Table II, and bi-serial r's were computed for each. The choice of bi-serial r as a statistic for this purpose depends upon its special properties. The important property for our present purpose is that, for normal distributions of the continuous (non-dichotomized) variate, the value of bi-serial r is invariant with respect to the point of dichotomy. Another property

which must be noted in passing is that bi-serial r has limits of ± 1 only when the continuous variate is normally distributed.

Table II gives the values of bi-serial r for a large number of points of dichotomy of the criterion. These values are also represented in Figure 1.

TABLE II

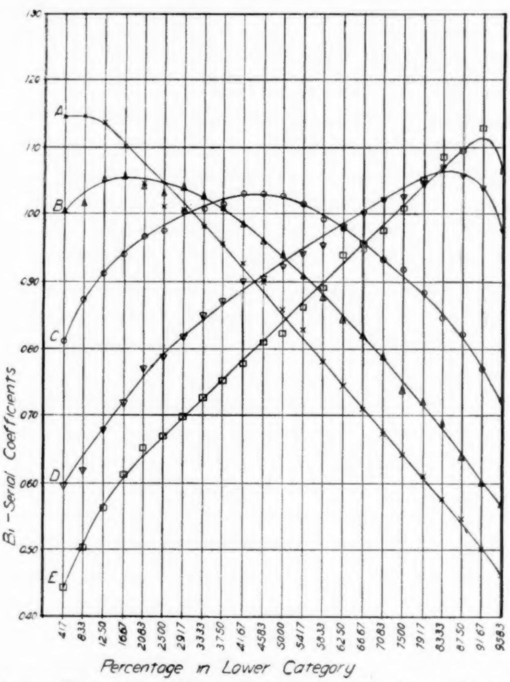
Bi-Serial Coefficients of Correlation of Subtests for Various Divisions of Criterion into Two Categories

Criterion		Coefficients				
Percentage in Upper Category	Percentage in Lower Category	A	B	C	D	E
4.17	95.83	.466	.570	.728	.976	1.066
8.33	91.67	.504	.600	.771	1.040	1.129
12.50	87.50	.549	.663	.823	1.059	1.096
16.67	83.33	.580	.689	.848	1.072	1.089
20.83	79.17	.611	.721	.884	1.046	1.051
25.00	75.00	.645	.744	.918	1.026	1.012
29.17	70.83	.675	.787	.934	1.022	.978
33.33	66.67	.711	.821	.949	1.001	.956
37.50	62.50	.746	.843	.977	.983	.940
41.67	58.33	.783	.875	.993	.953	.892
45.83	54.17	.826	.907	1.015	.941	.862
50.00	50.00	.860	.939	1.027	.924	.825
54.17	45.83	.900	.960	1.030	.904	.811
58.33	41.67	.927	.986	1.030	.900	.779
62.50	37.50	.957	1.012	1.015	.871	.752
66.67	33.33	.984	1.027	1.008	.850	.728
70.83	29.17	1.007	1.040	1.003	.818	.699
75.00	25.00	1.013	1.032	.978	.786	.670
79.17	20.83	1.048	1.042	.969	.768	.653
83.33	16.67	1.102	1.056	.940	.720	.613
87.50	12.50	1.139	1.054	.912	.679	.562
91.67	8.33	1.148	1.017	.871	.644	.503
95.83	4.17	1.148	1.007	.813	.597	.445

It is clearly seen from Figure 1 that Test A (the easy test) is most effective in separating off a small percentage from the lower part of the criterion distribution. Test E (difficult) is likewise most valid for separating off a small percentage of the criterion at the higher part of the distribution. The other tests have points or regions of maximum validity which are intermediate in the appropriate order. Test A is most valid for prediction to a two-categoried criterion which is divided at approximately -1.4σ. Test B is most valid for a criterion divided at approximately -1.0σ. Test C is most valid for a

criterion divided at —0.10σ. Test D is most valid for a criterion divided at approximately 1.0σ. Test E is most valid for a criterion divided at approximately +1.4σ.

These results prove, in a qualitative fashion, that fifty-item tests have pronounced differential validities appropriate to the various levels of difficulty. The objective of the next section is to outline in a



Bi-Serial Coefficients of Correlation for Various Dichotomies of the Criterion

FIGURE 1

more useful fashion a method of describing the difficulty of a test element and its validity in terms of a test discrimination function.

III. The Discrimination Function

1. The Theoretical Curve

The difficulty of a test element which is objective may be simply described by the percentage of the population failing (or passing) the

item, i.e., by q or p. This definition of difficulty makes the one measure (p) primarily dependent upon the distribution of abilities of the arbitrary population. Obviously, an element or task which can be correctly performed by p_1 per cent of a given population may be correctly performed by a higher percentage p_2 of another population with a higher mean. This dependence of a measure of difficulty on the parameters of the distribution function of the population studied may not be escaped; difficulty is merely an obverse of achievement. When we say that a mental task is very difficult, for instance, we are merely expressing the fact that few individuals, of some group we have in mind, are capable of performing the task. If we use p (the percentage of passes) as a measure of difficulty of a test item, we understand that p is an average value; in fact p may be regarded as the average score of the group upon a test consisting of that one item. If we have a single task which has positive validity for the prediction of some criterion, and which may be unequivocally evaluated in terms of success or failure, we should have for each successively greater criterion value c_i a corresponding greater percentage of passes p_i. The relation between p_i and c_i could be regarded as a discrimination function for the single item.

However, there are various reasons for not using the single item for the present purpose. One reason is that a single item is subject to great fluctuations in response, i.e., it is unreliable. In the second place, the validity of most items is so low that item curves are comparatively flat, and unrepresentative of the test discrimination function. In the third place, the p value (average difficulty) of the single item is fixed and cannot be systematically varied. The discussion of the device adopted in the subsequent treatment will make clear the significance of this objection. In so far as a single objective item makes a prediction of the criterion, the distribution has a point character. The record on the item, taken at its face value, makes a total of pN times qN unit discriminations, where N is the population. In other words pN individuals are judged as better than qN other individuals. The item behaves as if pqN^2 all-or-none judgments were made upon the abilities of individuals of the group.*

Let us consider a test element to be sufficiently well represented by a limited number of items, homogeneous as to content and number

*Obviously, the number of possible judgments on N individuals with respect to all others is $\frac{1}{2}N(N-1)$. Not all judgments are made by one item, and the maximum number of judgments is $\frac{1}{4}N^2$ which occurs when $p=q=.5$. If we assume that item validity varies directly with the number of point discriminations made by the item, then the item of 50 per cent difficulty is the most valid, other things being equal.

Cf. Thurstone, Thelma Gwinn. "The Difficulty of a Test and Its Diagnostic Value," *Journal of Educational Psychology*, 1932, **23**, 335-43.

passing each. Furthermore, for convenience let the items each be passed by approximately 50 per cent of the population. These requirements are approximately satisfied by the fifty item test which has been designated as Test C. Although the test is to be considered as a unit, the constants of the distribution of scores on Test C and of its bivariate distribution against the criterion were computed for use in the subsequent analysis. The difficulty of this test element is now systematically varied by setting standards of difficulty at different points in the distribution of scores on this fifty-item test. This makes possible, for each argument of difficulty, the description of the discrimination function in terms of parameters which have meaning in themselves, and in terms of the more generalized psychophysical theory. In Figure 2 the bivariate distribution of the test element

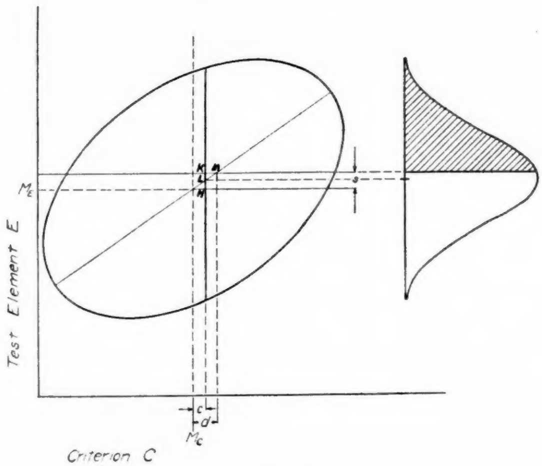


FIGURE 2

E and of the criterion C is represented in hypothetical terms. The means are M_e and M_c respectively. As a first approximation we assume linear regression, normal and homoscedastic E-arrays. Inspection of the obtained distribution suggests that this assumption is not unreasonable especially for values of E not too far from its mean. Let s, measured in standard deviation units from the mean of E, be the standard set for the difficulty of the test element, and constituting the operational determiner of difficulty of the element. One can express this difficulty of the element in terms of p, the percentage passing, but this is not immediately useful. This latter form of

expression becomes possible by reason of the fact that element E is dichotomized at s.

In Figure 2, the vertical line which contains the points K, L, and H represents any array in the bivariate distribution. The distribution of scores on the test element in this array is represented by the normal curve shown in Figure 2. In this array, the proportion of the group, all having a criterion score of c who exceed the standard s is indicated by the cross hatched portion of the curve at the right. This proportion varies with c, and will now be expressed as a function of c and the necessary parameters.

Let the deviation score in the element be designated by e.

Then the value of e corresponding to each c is given by the regression equation

$$e = b_{ec}\, c \,, \tag{1}$$

where b_{ec} is the coefficient of regression of e on c. For the array shown in Figure 2, e is the distance LH. Since L denotes the location of the mean of the distribution of the array, the array score which just equals the standard is represented by KL. This distance we shall call a. The expression for the value of the standard in the array is then

$$a = \frac{s - b_{ec}\, c}{\sigma_a} \,, \tag{2}$$

where σ_a is the standard deviation of the array, and is assumed to be constant for all arrays.

There will exist one array in which $a = 0$. This will occur when $s = b_{ec}\, c$. In this array, which can be represented by a vertical line passing through the point of intersection of the line of regression with the standard, the median array score just meets the standard. This point is M in Figure 2. We will now let the criterion score corresponding to this array be represented by

$$d = \frac{s}{b_{ec}} \,. \tag{3}$$

Then d is the parameter to be used in describing the difficulty of test element E when the standard is s. *The difficulty of a test element is defined as the standard score of the criterion measure at which the element (as a whole) is equally often passed and failed.*

Equation (2) may now be written, using the definition (3), thus:

$$a = \frac{b_{ec}(d - c)}{\sigma_a} \,. \tag{4}$$

According to the hypothesis adopted there will exist for each array, a theoretical proportion meeting or exceeding the standard of difficulty. The observed proportions in the twenty-four arrays will not agree exactly with the theoretical proportions. We will assume, consistently with the theory thus far applied to the problem, that the integral of the normal probability curve may be used as the theoretical curve. Clearly for a normal bivariate distribution, the proportions in the successive arrays vary with the corresponding criterion values exactly according to this function.*

If in equation (4) we now substitute $\sigma_d = \frac{\sigma_a}{b_{ec}}$, we have

$$a = \frac{d - c}{\sigma_d}. \tag{5}$$

The theoretical proportions are represented by

$$p = \frac{1}{N} \int_{\frac{d-c}{\sigma_d}}^{\infty} y \, dc , \tag{6}$$

where

$$y = \frac{N}{\sigma_d \sqrt{2\pi}} e^{-\frac{(d-c)^2}{2\sigma_d^2}}.$$

Following the usual procedure of minimizing the sum of the squares of the errors in the x (or base line) values corresponding to the proportions, we have as the expression to be minimized

$$v = \Sigma \left(\frac{d - c}{\sigma_d} - x \right)^2, \tag{7}$$

the summation to extend over all arguments of the criterion for which data are available. The term $\frac{d - c}{\sigma_d}$ corresponds to the theoretical, and the x to the observed proportions. The formal problem is now to find values of d and σ_d which will make this sum of squares a minimum.

Differentiating equation (7) with respect to $1/\sigma_d$ and d/σ_d in turn, we have as normal equations

*The assumption of normality and homoscedasticity of arrays is tantamount to the choice of the integral of the normal curve as the theoretical function. The curve fitting procedure will hereafter be described in outline only, since it is adequately treated in other applications, especially in the literature on the Constant Method in psychophysics.

See Kelley, Truman L. Statistical Method. New York: Macmillan Co., 1924. Pp. 326-30.

$$\frac{1}{\sigma_d}\sum w\,c - \frac{d}{\sigma_d}\sum w - \sum w\,x = 0\ , \tag{8}$$

$$\frac{1}{\sigma_d}\sum w\,c^2 - \frac{d}{\sigma_d}\sum w\,c - \sum w\,x\,c = 0\ . \tag{9}$$

The w's in the equations are the usual weights used in the Constant Method.

Solving (8) and (9) simultaneously,

$$\sigma_d = \frac{\sum w \cdot \sum wc^2 - (\sum wc)^2}{\sum w \cdot \sum wxc - \sum wx \cdot \sum wc}\ , \tag{10}$$

$$d = \frac{\sum wc \cdot \sum wxc - \sum wx \cdot \sum wc^2}{\sum w \cdot \sum wxc - \sum wx \cdot \sum wc}\ . \tag{11}$$

Five degrees of difficulty were arbitrarily created by setting up various standards for passing the test element. These standards, and the computed values of σ_d and d, are presented in Table III.

TABLE III.

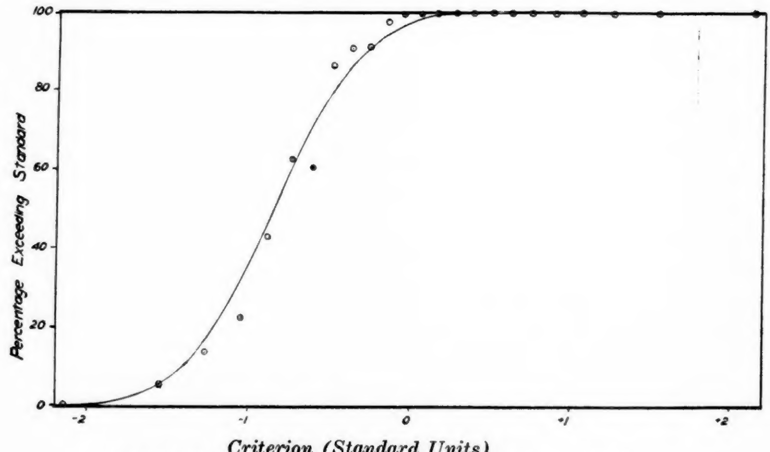
Parameters of Discrimination Function for Various Standards Subtest C

Standard in score Units of the Element	Standard in Standard Scores	Percentage of Passes p	Index of Validity σ_d	Liminal Difficulty d
37.07	1.0	22.0	0.550	0.92
31.59	0.5	34.6	0.412	0.34
26.11	0.0	51.8	0.316	–0.05
20.63	–0.5	64.9	0.317	–0.39
15.14	–1.0	78.2	0.433	–.085

This table shows the fairly close relation between the difficulty of the test element and the standard adopted for passing the element.

In Figures 3 to 7 the observed proportions of each of the twenty-four criterion intervals who "pass" the element for the various standards of difficulty are shown, with the theoretical curve fitted to each.

The test element, with standard set to allow 51.8 per cent of passes, has a difficulty measure of —0.05, measured in the standard units of the criterion. Under these conditions the test element is most valid for discriminating between individuals at or near the mean of the criterion group. The element does not differentiate between the abilities of individuals in the highest 20 per cent of the criterion group. The same is true of the lowest 20 per cent of the group. Considering that the region between the inflection points of the under-

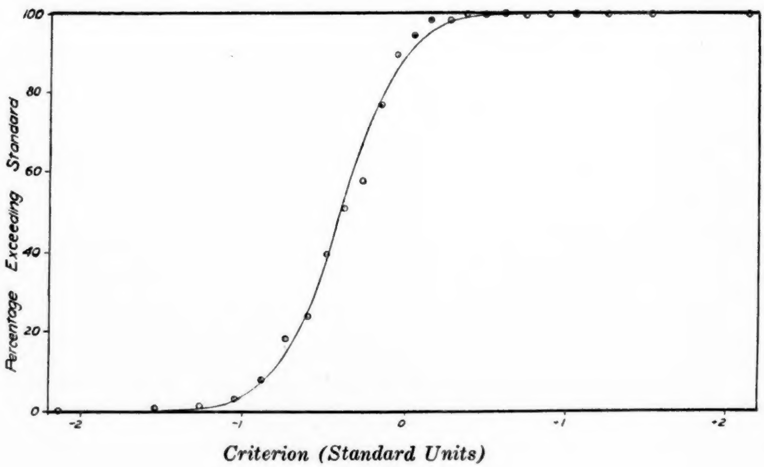


Criterion (Standard Units)

Discrimination Function for Test Element *C*

FIGURE 3

Standard = $+1.0\sigma$



Criterion (Standard Units)

FIGURE 4

Discrimination Function for Test Element *C*

Standard: = $+0.5\sigma$

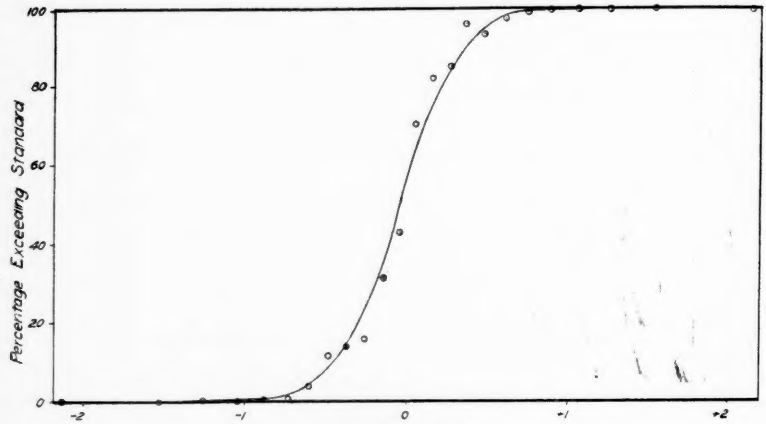


FIGURE 5
Discrimination Function for Test Element C
Standard: Mean $= 0.0\sigma$

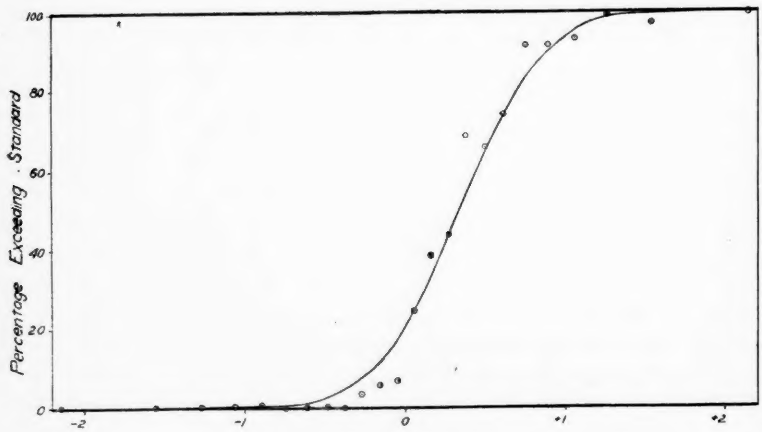
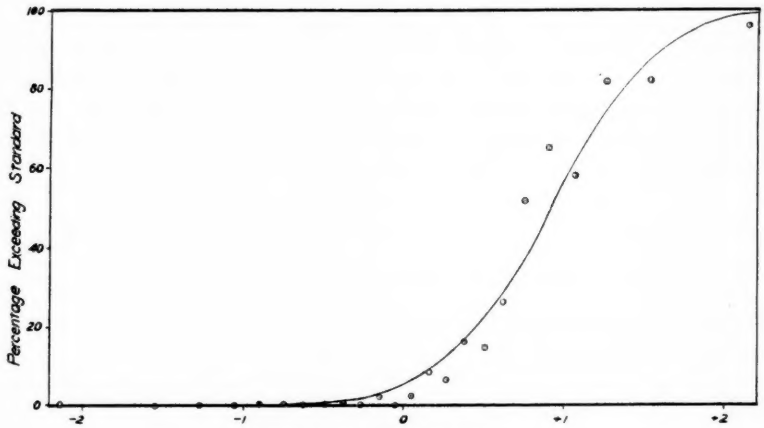


FIGURE 6
Discrimination Function for Test Element C
Standard: $= -1.0\sigma$



Discrimination Function for Test Element *C*
FIGURE 7

Standard: $= -0.5\sigma$

lying normal curve is to be regarded as the practical limit of the element's effectiveness, it can be said that a high validity element of this degree of difficulty is functional only between criterion scores of -0.37σ to 0.27σ. This is equivalent to saying that the element thus defined has validity for approximately the middle 25 per cent of the criterion group.

When the difficulty of the test element is raised so that the liminal difficulty is 0.34 (percentage of passes is 34.6), the area of effectiveness is in the direction of higher criterion scores. The region between one standard deviation above and below the liminal difficulty is defined by the limits 0.07σ and 0.75σ, which includes approximately 25 per cent of the criterion group. The measure of precision is the standard deviation of the theoretical curve. Its value is 0.412 in this case. When the standard of difficulty of the element is set at a still higher point, viz., 37.07 or 1σ in the element variable, the element is functional in discriminating between individuals only in the upper half of the criterion group. This element has a difficulty such that practically none of those in the lower half of the criterion group are able to perform the task set. The greater σ_d indicates that it is a less valid measuring instrument for the purpose of separating those with criterion scores above 1σ from those with scores below 1σ, than is the same element for standards of difficulty which will divide the crite-

rion group at points nearer to the mean. Likewise, standards of difficulty set for this element at points below the mean result in specific validity for a restricted sub-range of criterion ability, and practically no validity for other sub-ranges of ability. Likewise the element predicts the criterion less well in general as the standard departs from a liminal difficulty of 0σ.

IV. Relation to Coefficient of Validity

In order to simplify equation (4), the substitution of σ_d for $\frac{\sigma_a}{b_{ec}}$ was made. This substitution was not arbitrary. We may write from ordinary correlation theory,

$$\sigma_a = \sqrt{1 - r^2_{ec}} \cdot \sigma_e \, , \tag{12}$$

and

$$b_{ec} = r_{ec} \frac{\sigma_e}{\sigma_c} \, . \tag{13}$$

Then we may rewrite (5) by substitution of (12) and (13), and simplify it to

$$\sigma_d = \frac{\sqrt{1 - r^2_{ec}}}{r_{ec}} \, , \tag{14}$$

when $\sigma_c = 1$. When (14) is solved for r_{ec}, we have

$$r_{ec} = \sqrt{\frac{1}{1 + \sigma_d^2}} \, . \tag{15}$$

Equation (15) gives us a method of estimating the validity coefficient. In Table IV is presented the validity coefficients, as estimated by equation (15), and as directly computed.

TABLE IV

Comparison of Measures of Validity of Elements of Varying Difficulty

Element	Validity Coefficient: (1) As Estimated from the Discrimination Function	Validity Coefficient: (2) As Computed Directly	Standard Error of (2)
C; Standard: 1.0σ	0.876	0.896	0.0057
C; Standard: 0.5σ	0.924	0.931	0.0038
C; Standard: $0.\ \sigma$	0.954	0.958	0.0024
C; Standard: -0.5σ	0.953	0.954	0.0026
C; Standard: -1.0σ	0.918	0.941	0.0033
D; Standard: Median	0.892	0.902	0.0054

It is seen from the table that the validity of a test element as estimated from the dispersion parameter of the discrimination curve is a fairly close approximation to the validity as computed in correlational terms. This indicates that the assumptions back of the rationale were fairly justified. Moreover, in situations where such validity coefficients are not available, they may be easily estimated from the precision parameter of the psychometric curve which we have called the discrimination function.

V. DISCUSSION AND CONCLUSIONS

1. Although not directed at precisely the same problem, the present investigation confirms the conclusion of Thelma Gwinn Thurstone that a test composed of items of 50 per cent difficulty has a general validity which is higher than tests composed of items of any other degree of difficulty.*

Her results, however, are not inconsistent with the hypothesis that a test or test element of a difficulty other than the optimum might have a satisfactory differential validity which is confined to specific intervals of the criterion measure. Thus a test or test element, while indubitably having only fair general or average validity, might conceivably be highly valid for differentiating between individuals on some specified part of the criterion measure. The maximum general validity of elements of 50 per cent difficulty might be easily interpreted in terms of a differential validity peculiar to each degree of difficulty. If we assume that each criterion score c can be best distinguished from adjacent scores by an element of difficulty d, each value of c being associated with an element of optimal difficulty d_c, then it is clear that when $d = 0$ ($p = 50$ per cent), the condition of maximum general validity is attained, since $d = 0$ is the best representative of a distribution $(d_1, d_2, \cdots d_n)$ of different optimal difficulties for the various values of c. Any other value of d would produce a greater average squared deviation from the respective optima for the various criterion scores.

2. It is definitely established by these experiments that tests of different difficulty will predict to a two-categoried criterion with different degrees of effectiveness. If it is desired to separate off a minor proportion from the lower end of the distribution of criterion scores, then an easy test has much greater validity than have more difficult tests. Moreover, the smaller the minor proportion to be separated from the criterion group at the lower end, the easier should the

**Op. cit.*

test be. The converse situation applies to minor proportions to be marked off from the upper end of the criterion group. It is to be noted, also, that if the general validity of a test can be regarded as a sort of average of the various differential validities, then its general or average validity decreases as the division point for the maximal prediction into two categories departs from 0σ, the mean of the distribution of criterion scores.

3. The terms "differential validity" and "liminal difficulty" have been repeatedly used in this investigation. The term "differential validity" refers to the capacity of the measuring instrument in discriminating between various levels of ability. When the differential validity of a measuring device varies from one region or interval of ability to another, it becomes pertinent to inquire at what specific point this discriminative capacity is a maximum. The criterion score at this maximum we have defined as difficulty, or liminal difficulty. The term "liminal" is aptly used because of its close analogy to the limen of sensitivity found from psycho-physical experiments. The method of describing the test element as a discrimination function has at least two advantages. One advantage is that the same type of mathematical function can be used as in the conventional psychophysical theory; a skew function can be easily adopted if the use of a third parameter is desirable by reason of small elements of extreme positive or negative degrees of difficulty. A more important advantage is that the parameters of the curve have meaning built into them by the manner of derivation. In fact, the meaning of the parameters as used in the test setting have been expressed, to a fair approximation, in terms of the customary measures of validity and difficulty.

A further practical advantage of the method of measuring and expressing difficulty is that graphic methods are available. With the use of probability paper, the liminal difficulty of a test element and its validity may be quickly obtained.

4. The procedures outlined in the foregoing definitely point to the unsatisfactory nature of the common practice in the construction of tests of letting difficulty take care of itself. It is true that a test, no matter how the difficulty of its elements is distributed, will give a distribution of scores which may have practical usefulness. We cannot assume, however, unless the test is exceedingly long, that a chance distribution of difficulty will give scores which are linear with true measures of the ability. If, on the other hand, we have some specific purpose of prediction in mind, we may utilize the differential validities of test elements to our advantage. Suppose, for example,

that a test of clerical aptitude is meant to sort out the best 15 per cent of all applicants. This is on the assumption that the labor market is such that one hundred persons will apply for fifteen positions. It is then clear that the optimal difficulty of test elements should be in the neighborhood of $+1\sigma$ and that easier tasks would give us discriminations between individuals in whom we are not interested. The proper choice of difficulty gives us the maximum discrimination between the applicants we care to consider seriously. The converse consideration would apply to any situation in which a minor proportion from the lower end of a distribution is to be separated off for any purpose. Under any circumstances involving educational or psychological measurement, the distribution of difficulty of the elements or tasks can be arranged to fulfill more accurately the purposes of the measurement.

PSYCHOMETRIKA—VOL. 1, NO. 2
JUNE, 1936

NOTE ON COMPUTATION OF BI-SERIAL CORRELATIONS IN ITEM EVALUATION

JACK W. DUNLAP

Fordham University, New York

By the use of an algebraic variant of the ordinary formula for bi-serial correlation, tables, and graphic devices, a time-saving systematic procedure for the computation of bi-serial correlation coefficients is outlined for application to the evaluation of items of a test. A table of $\frac{p}{z}$ for arguments of $p = .000$ to $p = .999$ is given.

One of the most laborious steps in test construction is that of item validation. Various techniques have been presented for validating items, one of which, the method of bi-serial correlation has attained considerable favor. A practical disadvantage of this method is that it is time consuming. The purpose of this paper is to outline a method which materially reduces the labor of computation whenever it is necessary to compute many such correlation coefficients.

The formula for the bi-serial correlation is usually written

$$r_{bi} = \frac{M_P - M_F}{\sigma} \cdot \frac{pq}{z} , \tag{1}$$

but can be rewritten in the form

$$r_{bi} = \frac{M_P - M_{T \cdot D}}{\sigma_{T \cdot D}} \cdot \frac{p}{z} , \tag{2}$$

where

$M_{T \cdot D}$ is the mean of the criterion scores for the total group;

M_P is the mean of the criterion scores for the group passing the item being studied;

$\sigma_{T \cdot D}$ is the standard deviation of the criterion scores for the total group;

z is the ordinate corresponding to p.

In formula (1) it is necessary to compute the mean criterion scores for those who pass the item and for those who fail the item. Therefore if we are determining validity coefficients for 500 items (which are scored right or wrong) it is necessary to compute 1000 means. If the second formula is used it is necessary to determine

only the mean score of all the subjects, and then for each item the mean criterion score of the subjects who succeed with the item. Thus, the computation of 500 means is eliminated.

The following steps have been found economical in securing the necessary data for computing bi-serial correlations when Hollerith Tabulating equipment is not available.

1. On each test blank write the criterion score. This may be either the total score on the test of which the item is a part or a score from an independent criterion.

2. The criterion scores should be expressed as T scores, thus simplifying the formula for r to

$$r_{bi} = \frac{M_1 - 50.0}{10} \cdot \frac{p}{z} .$$

The subtraction $M_1 - 50.0$ may be done mentally and the division consists simply of moving the decimal point one place to the left. Raw scores can be readily changed to T scores by a line graph which can easily be constructed.

(*a*). In order to construct a line graph to change raw scores to T scores, we may write

$$T = 50 + \frac{10(X - M)}{\sigma}$$

Rewriting to get the formula in a convenient linear form gives

$$T = 50 + (10/\sigma)X - (10/\sigma)M \qquad \text{or}$$

$$T = a + bX - k \qquad \text{or}$$

$$T = A + bX .$$

(*b*). Suppose in a given problem that $M = 82.64$ and $\sigma = 12.23$, then

$$T = 50 + (10/12.23)X - (10/12.23)(82.64) \qquad \text{or}$$

$$T = .8177X - 20.52 .$$

(*c*). Now take a sheet of graph paper and plot the equation. Paper ruled ten to the inch is quite convenient. Care should be taken to plot the points accurately and the line should be drawn in with a fine nibbed pen. It is convenient to plot T on the Y axis or ordinate and X on the abscissa.

As the equation is linear three points are sufficient to determine and verify the line. Two of the three points can be readily determined, for when X is set equal to zero, the T value is A, the constant, and when X is set equal to the mean, then T equals 50.

(*d*). After the line of the equation has been drawn, project the values written on the T and X axes to the line and label. (See Fig. 1.) The T values will appear on one

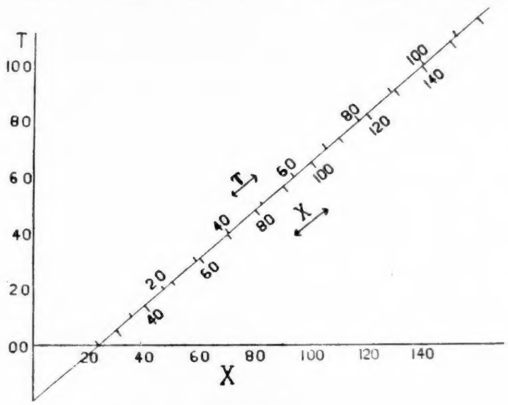


FIGURE 1

side of the line and the X values on the other, so that for a given value of X, T can be read directly.

As a final check on the accuracy of the line graph determine two or three values by the graph and verify by calculation. If the chart is drawn reasonably carefully the result will be read correctly to the first two places with a negligible error in the first decimal place.

(*e*). Using the line graph, express the criterion score for each blank as a T score.

3. Sort the blanks into piles, each pile corresponding to a class interval. Suppose the range of criterion scores is 74, i.e., from 12 to 86, so that grouping by three's we would have 25 piles. The first pile should contain all papers whose criterion score falls between 12 to 14.

4. Consider question No. 1. Go through the papers in the first pile. Here the class interval is 12-14. Count the number

who have correctly marked the question. Record this number on the basic data sheet in column one opposite the class interval 12-14. (See the basic data sheet.)

Now go through this pile of papers again and determine the number of papers on which question No. 2 was correctly marked. Record this number on the basic data sheet in column 2 opposite the interval 12-14. In a similar manner determine the number of correct responses for each of the remaining items and record in the appropriate column on the basic data sheet.

5. Next analyze the papers in the pile 15-17 to determine the number of correct responses for each item. These values are placed in row two of the basic data sheet. Similarly analyze and record the number of successes for each of the several piles. When this is completed the data are available for computing the mean criterion score for the individuals who passed each item.

6. The computation of the means can be expedited considerably by computing the sum of T or X' values for two questions at a time. This is done by setting the two class frequencies in the keyboard of a calculating machine and multiplying by the X' value assigned to that class.

7. Record $\sum X'$ for the items in the designated blanks at the bottom of the page. Also record the value N_p for each item, the number passing the item.

8. Compute all the means keeping the results to one decimal place. Each of these means is the mean score for the subjects who passed a particular item. The value M_p may be recorded in the line M_P, or what is more advisable do not record M_P, but immediately subtract 50 and move the decimal point one place to the left. The result will be $(M_P - M_T)/\sigma$, or specifically, as T scores are being used, $(M_P - 50)/10$, and should be recorded in the space labeled $(M_p - M_T)/\sigma$. Hereafter this expression will be referred to as Δ. Note: In all cases where the item has a positive validity the value of M_P will exceed 50.0. Items which have a mean value less than 50.0 have negative validity and probably should be discarded, unless examination of the item shows that the scoring key is in error.

9. Prepare a table or line graph giving the percentages for all values of N_P/N where N_P is the number passing an item and

N is the total population. In case the percentage of successes for an item is less than 5 or greater than 95 the item probably should be discarded due to the unreliability of these percentages.

10. Now by means of Table I determine the p/z value for each item. This may or may not be written on the basic data sheet at the option of the investigator.
11. Set this value, p/z, in a calculating machine and multiply by the corresponding Δ value. The result is the desired bi-serial correlation and should be recorded.
12. If it is desirable to weight the items, serviceable weights can be had by using the r values, first multiplying by ten to eliminate the decimals.
13. A nomograph for determining r when p and $(M_1 - M_T)/\sigma$ or Δ is known, is given by the writer in this issue. This nomograph is sufficiently accurate for practically all purposes.

In an investigation under the writer's direction it was necessary to compute four bi-serial correlations (one for each of four possible responses to the item) for each of 870 items. The population used was 200. Thus it was necessary to compute 3480 correlations. The entire job was completed in approximately 250 working hours, a rate of approximately 14 correlations an hour.

TABLE I

TABLE OF p and p/z COMPUTED FROM THE KELLEY-WOOD TABLE*

P	000	001	002	003	004	005	006	007	008	009	P
.00	—	.2970	.3155	.3279	.3376	.3458	.3529	.3592	.3650	.3702	.00
.01	.3752	.3798	.3842	.3883	.3923	.3961	.3997	.4032	.4066	.4099	.01
.02	.4131	.4161	.4192	.4221	.4250	.4278	.4305	.4332	.4358	.4384	.02
.03	.4409	.4434	.4458	.4482	.4506	.4530	.4553	.4575	.4598	.4620	.03
.04	.4642	.4663	.4685	.4706	.4727	.4747	.4768	.4788	.4808	.4828	.04
.05	.4848	.4868	.4887	.4906	.4925	.4944	.4963	.4982	.5000	.5019	.05
.06	.5037	.5055	.5073	.5091	.5109	.5126	.5144	.5162	.5179	.5196	.06
.07	.5213	.5231	.5248	.5265	.5282	.5298	.5315	.5332	.5348	.5365	.07
.08	.5382	.5398	.5414	.5430	.5446	.5462	.5479	.5495	.5511	.5526	.08
.09	.5542	.5558	.5574	.5589	.5605	.5621	.5636	.5652	.5667	.5683	.09
.10	.5698	.5713	.5729	.5744	.5759	.5774	.5790	.5805	.5820	.5835	.10
.11	.5850	.5865	.5880	.5895	.5910	.5925	.5940	.5954	.5969	.5984	.11
.12	.5999	.6014	.6028	.6043	.6058	.6072	.6087	.6102	.6116	.6131	.12
.13	.6145	.6160	.6174	.6189	.6203	.6218	.6232	.6247	.6261	.6276	.13
.14	.6290	.6304	.6319	.6333	.6347	.6362	.6376	.6390	.6405	.6419	.14
.15	.6433	.6448	.6462	.6476	.6491	.6505	.6519	.6533	.6547	.6562	.15
.16	.6576	.6590	.6604	.6619	.6633	.6647	.6661	.6675	.6690	.6704	.16
.17	.6718	.6732	.6746	.6761	.6775	.6789	.6803	.6818	.6831	.6846	.17
.18	.6860	.6874	.6888	.6902	.6916	.6931	.6945	.6959	.6973	.6987	.18
.19	.7002	.7016	.7030	.7044	.7058	.7073	.7087	.7101	.7115	.7130	.19
.20	.7144	.7158	.7172	.7187	.7201	.7215	.7229	.7244	.7258	.7272	.20
.21	.7287	.7301	.7315	.7330	.7344	.7358	.7373	.7387	.7401	.7416	.21
.22	.7430	.7444	.7459	.7473	.7488	.7502	.7517	.7531	.7546	.7560	.22
.23	.7575	.7589	.7604	.7618	.7633	.7647	.7662	.7676	.7691	.7706	.23
.24	.7720	.7735	.7749	.7764	.7779	.7794	.7808	.7823	.7838	.7852	.24
.25	.7867	.7882	.7897	.7912	.7926	.7941	.7956	.7971	.7986	.8001	.25
.26	.8016	.8031	.8046	.8061	.8076	.8091	.8106	.8121	.8136	.8151	.26
.27	.8166	.8181	.8196	.8211	.8226	.8242	.8257	.8272	.8287	.8303	.27
.28	.8318	.8333	.8349	.8364	.8379	.8395	.8410	.8426	.8441	.8457	.28
.29	.8472	.8488	.8503	.8519	.8534	.8550	.8566	.8581	.8597	.8613	.29
.30	.8628	.8644	.8660	.8676	.8691	.8707	.8723	.8739	.8755	.8771	.30
.31	.8787	.8803	.8819	.8835	.8851	.8867	.8883	.8900	.8916	.8932	.31
.32	.8948	.8965	.8981	.8997	.9014	.9030	.9046	.9063	.9079	.9096	.32
.33	.9112	.9129	.9145	.9162	.9179	.9195	.9212	.9229	.9246	.9262	.33
.34	.9279	.9296	.9313	.9330	.9347	.9364	.9381	.9398	.9415	.9432	.34
.35	.9449	.9466	.9484	.9501	.9518	.9536	.9553	.9570	.9588	.9605	.35
.36	.9623	.9640	.9658	.9675	.9693	.9711	.9728	.9746	.9764	.9782	.36
.37	.9800	.9817	.9835	.9853	.9871	.9889	.9907	.9926	.9944	.9962	.37
.38	.9980	.9998	1.0017	1.0035	1.0053	1.0072	1.0090	1.0109	1.0128	1.0146	.38
.39	1.0165	1.0183	1.0202	1.0221	1.0240	1.0259	1.0277	1.0296	1.0315	1.0334	.39
.40	1.0353	1.0373	1.0392	1.0411	1.0430	1.0450	1.0469	1.0488	1.0508	1.0527	.40
.41	1.0547	1.0566	1.0586	1.0606	1.0625	1.0645	1.0665	1.0685	1.0705	1.0725	.41
.42	1.0745	1.0765	1.0785	1.0805	1.0825	1.0845	1.0866	1.0886	1.0906	1.0927	.42
.43	1.0948	1.0968	1.0989	1.1009	1.1030	1.1051	1.1072	1.1093	1.1113	1.1134	.43
.44	1.1156	1.1177	1.1198	1.1219	1.1240	1.1262	1.1283	1.1305	1.1326	1.1348	.44
.45	1.1369	1.1391	1.1413	1.1434	1.1456	1.1478	1.1500	1.1522	1.1544	1.1567	.45
.46	1.1589	1.1611	1.1633	1.1656	1.1678	1.1701	1.1723	1.1746	1.1769	1.1792	.46
.47	1.1815	1.1838	1.1861	1.1884	1.1907	1.1930	1.1953	1.1976	1.2000	1.2023	.47
.48	1.2047	1.2071	1.2094	1.2118	1.2142	1.2166	1.2190	1.2214	1.2238	1.2262	.48
.49	1.2286	1.2311	1.2335	1.2360	1.2384	1.2409	1.2433	1.2458	1.2483	1.2508	.49

*"The values in the table were computed by dividing the value for p by the value for z found in the Kelley-Wood Table in Kelley, T. L., Statistical Method. The table was checked by differencing and again by a complete recomputation of the entire series."

TABLE OF p AND p/z COMPUTED FROM THE KELLEY-WOOD TABLE

Page two

P	000	001	002	003	004	005	006	007	008	009	P
.50	1.2533	1.2558	1.2583	1.2609	1.2634	1.2659	1.2685	1.2711	1.2736	1.2762	.50
.51	1.2788	1.2814	1.2840	1.2866	1.2892	1.2918	1.2945	1.2971	1.2998	1.3024	.51
.52	1.3051	1.3078	1.3105	1.3131	1.3159	1.3186	1.3213	1.3240	1.3268	1.3295	.52
.53	1.3323	1.3351	1.3378	1.3406	1.3434	1.3462	1.3490	1.3519	1.3547	1.3576	.53
.54	1.3604	1.3633	1.3662	1.3691	1.3720	1.3749	1.3778	1.3807	1.3837	1.3866	.54
.55	1.3896	1.3925	1.3955	1.3985	1.4015	1.4045	1.4076	1.4106	1.4137	1.4167	.55
.56	1.4198	1.4229	1.4260	1.4291	1.4322	1.4353	1.4385	1.4416	1.4448	1.4480	.56
.57	1.4512	1.4544	1.4576	1.4608	1.4641	1.4673	1.4706	1.4739	1.4772	1.4805	.57
.58	1.4838	1.4871	1.4905	1.4938	1.4972	1.5006	1.5040	1.5074	1.5108	1.5142	.58
.59	1.5177	1.5212	1.5246	1.5281	1.5317	1.5352	1.5387	1.5423	1.5458	1.5494	.59
.60	1.5530	1.5566	1.5603	1.5639	1.5676	1.5713	1.5749	1.5786	1.5824	1.5861	.60
.61	1.5899	1.5936	1.5974	1.6012	1.6051	1.6089	1.6128	1.6166	1.6205	1.6244	.61
.62	1.6283	1.6323	1.6362	1.6402	1.6442	1.6482	1.6523	1.6563	1.6604	1.6645	.62
.63	1.6686	1.6727	1.6768	1.6810	1.6852	1.6894	1.6936	1.6978	1.7021	1.7064	.63
.64	1.7107	1.7150	1.7194	1.7237	1.7281	1.7325	1.7369	1.7414	1.7459	1.7503	.64
.65	1.7549	1.7594	1.7640	1.7685	1.7731	1.7778	1.7824	1.7871	1.7918	1.7965	.65
.66	1.8013	1.8060	1.8108	1.8156	1.8205	1.8253	1.8302	1.8352	1.8401	1.8451	.66
.67	1.8501	1.8551	1.8601	1.8652	1.8703	1.8754	1.8806	1.8858	1.8910	1.8962	.67
.68	1.9015	1.9068	1.9121	1.9175	1.9229	1.9283	1.9337	1.9392	1.9447	1.9503	.68
.69	1.9558	1.9614	1.9670	1.9727	1.9784	1.9841	1.9899	1.9957	2.0015	2.0074	.69
.70	2.0133	2.0192	2.0252	2.0312	2.0372	2.0433	2.0494	2.0555	2.0617	2.0679	.70
.71	2.0742	2.0805	2.0868	2.0932	2.0996	2.1060	2.1125	2.1191	2.1256	2.1322	.71
.72	2.1389	2.1456	2.1523	2.1591	2.1659	2.1728	2.1797	2.1867	2.1937	2.2007	.72
.73	2.2078	2.2149	2.2221	2.2294	2.2366	2.2440	2.2514	2.2588	2.2663	2.2738	.73
.74	2.2814	2.2890	2.2967	2.3044	2.3122	2.3201	2.3280	2.3359	2.3440	2.3520	.74
.75	2.3601	2.3683	2.3766	2.3849	2.3932	2.4017	2.4102	2.4187	2.4273	2.4360	.75
.76	2.4447	2.4535	2.4624	2.4713	2.4803	2.4894	2.4986	2.5078	2.5171	2.5264	.76
.77	2.5358	2.5453	2.5549	2.5646	2.5743	2.5841	2.5940	2.6039	2.6140	2.6241	.77
.78	2.6343	2.6446	2.6550	2.6654	2.6760	2.6866	2.6973	2.7081	2.7191	2.7300	.78
.79	2.7411	2.7523	2.7636	2.7750	2.7865	2.7981	2.8097	2.8215	2.8334	2.8454	.79
.80	2.8575	2.8698	2.8821	2.8945	2.9071	2.9197	2.9325	2.9454	2.9585	2.9716	.80
.81	2.9849	2.9983	3.0118	3.0255	3.0393	3.0532	3.0673	3.0815	3.0959	3.1104	.81
.82	3.1250	3.1398	3.1547	3.1698	3.1851	3.2005	3.2160	3.2318	3.2476	3.2637	.82
.83	3.2799	3.2964	3.3129	3.3297	3.3466	3.3638	3.3811	3.3986	3.4163	3.4342	.83
.84	3.4524	3.4707	3.4892	3.5080	3.5269	3.5461	3.5656	3.5852	3.6051	3.6252	.84
.85	3.6456	3.6662	3.6871	3.7082	3.7296	3.7513	3.7732	3.7954	3.8179	3.8407	.85
.86	3.8638	3.8872	3.9109	3.9350	3.9593	3.9840	4.0090	4.0344	4.0601	4.0862	.86
.87	4.1126	4.1394	4.1666	4.1942	4.2222	4.2506	4.2794	4.3087	4.3384	4.3685	.87
.88	4.3991	4.4302	4.4618	4.4938	4.5264	4.5595	4.5931	4.6272	4.6620	4.6973	.88
.89	4.7331	4.7697	4.8068	4.8445	4.8829	4.9220	4.9618	5.0023	5.0435	5.0855	.89
.90	5.1283	5.1718	5.2162	5.2614	5.3075	5.3545	5.4023	5.4512	5.5010	5.5519	.90
.91	5.6038	5.6567	5.7108	5.7661	5.8225	5.8801	5.9391	5.9993	6.0609	6.1239	.91
.92	6.1884	6.2543	6.3219	6.3911	6.4619	6.5346	6.6090	6.6853	6.7636	6.8440	.92
.93	6.9264	7.0112	7.0982	7.1876	7.2796	7.3742	7.4715	7.5717	7.6749	7.7813	.93
.94	7.8910	8.0042	8.1210	8.2416	8.3661	8.4950	8.6283	8.7663	8.9093	9.0574	.94
.95	9.2111	9.3708	9.5366	9.7089	9.8884	10.0752	10.2701	10.4733	10.6858	10.9078	.95
.96	11.1403	11.3840	11.6396	11.9083	12.1910	12.4887	12.8028	13.1350	13.4866	13.8597	.96
.97	14.2559	14.6779	15.1284	15.6100	16.1266	16.6824	17.2817	17.9299	18.6339	19.4007	.97
.98	20.2404	21.1632	22.1831	23.3159	24.5828	26.0100	27.6291	29.4847	31.6322	34.1506	.98
.99	37.1454	40.7718	45.2555	50.9570	58.4603	68.8105	84.0719	108.9737	157.4132	296.7033	.99

TABLE II

Basic data sheet for bi-serial correlations in item analysis

Study...

← item →

Class	X'	1	2	3	4	5	6	7	8	9	10	11	12
84-86	24												
81-83	23												
78-80	22												
75-77	21												
24-26	4	27	15	10									
21-25	3	32	6	9									
18-20	2	5	7	2									
15-17	1	4	2	0									
12-14	0	1	0	2									
$\Sigma X'$													
N_p													
M_p													
$(M_p - M_T)/\sigma$													
N_p/N													
p/z													
r													

Note that N, M_T and σ_T (σ) are constant for all items based on the total distribution.

PSYCHOMETRIKA—VOL. 1, NO. 2
JUNE, 1936

NOMOGRAPH FOR COMPUTING BI-SERIAL CORRELATIONS

JACK W. DUNLAP

Fordham University

The widespread use of bi-serial r in item analysis makes a quick and accurate method of evaluating the formula desirable. Such a method is provided by nomographs. Dunlap and Kurtz* in their Nomograph No. 47 give a method of solving the better known statement of the formula, namely

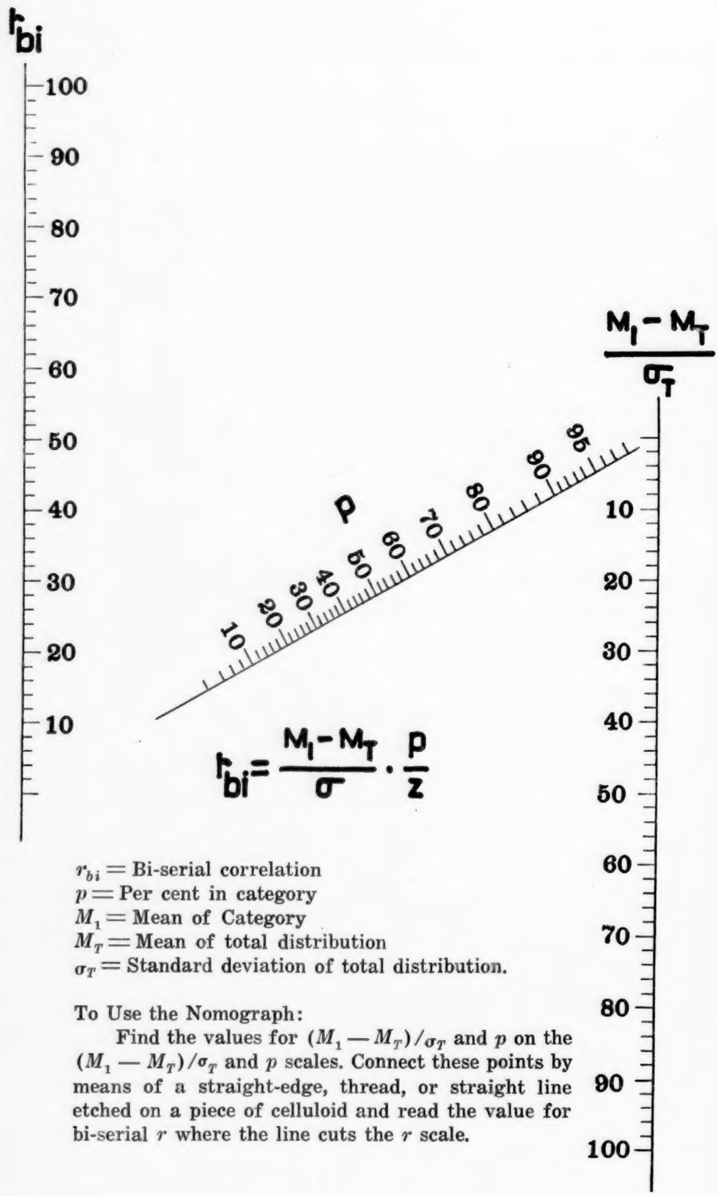
$$r_{bi} = \frac{M_p - M_q}{\sigma} \cdot \frac{pq}{z} . \tag{1}$$

As the writer has shown elsewhere in this Journal a form more convenient for solution in problems of item analysis is

$$r_{bi} = \frac{M_1 - M_T}{\sigma} \cdot \frac{p}{z} . \tag{2}$$

The nomograph for the solution of this expression together with the directions for using the nomograph are given on the next page.

*Dunlap, Jack W. and Kurtz, Albert K., Handbook of Statistical Nomographs, Tables, and Formulas. *World Book Co.* 1932, Yonkers, New York.



r_{bi} = Bi-serial correlation
p = Per cent in category
M_1 = Mean of Category
M_T = Mean of total distribution
σ_T = Standard deviation of total distribution.

To Use the Nomograph:

Find the values for $(M_1 - M_T)/\sigma_T$ and p on the $(M_1 - M_T)/\sigma_T$ and p scales. Connect these points by means of a straight-edge, thread, or straight line etched on a piece of celluloid and read the value for bi-serial r where the line cuts the r scale.

PSYCHOMETRIKA—VOL. 1, NO. 2
JUNE, 1936

LIST OF MEMBERS OF PSYCHOMETRIC SOCIETY

As of June 20, 1936

(Personal and institutional subscriptions to the journal are not included.)

Please notify the Secretary, Dr. Paul Horst, Personnel Research Department, The Procter and Gamble Company, Cincinnati, Ohio, U.S.A., of any change of address.

ACHILLES, DR. PAUL S., 522 5th Avenue, New York City.
ACKERSON, DR. LUTON, Diagnostic Depot, 2 Woodruff Road, Joliet, Illinois.
ADKINS, MISS DOROTHY, Ohio State University, Columbus, Ohio.
ALEXANDER, MR. W. P., 263 High Street, Walthamstow, E 17, England.
ALTIERE, MR. EDWARD S. A., 31 Woodman Street, Providence, Rhode Island.
ANASTASI, DR. ANNE, Barnard College, Columbia University, New York City.
ANDERSON, DR. L. DEWEY, Brush Foundation, Western Reserve University, Cleveland, Ohio.

BEERS, DR. F. S., Strohan House, University of Georgia, Athens, Georgia.
BENGE, MR. EUGENE J., Apt. 16, Twin Oaks Apts, 202 Twin Oaks Road, Akron, Ohio.
BINGHAM, DR. W. V., 29 West 39th St., New York City.
BODER, DR. DAVID, Lewis Institute, 1951 W. Madison St., Chicago, Ill.
BORING, DR. EDWIN G., Emerson Hall, Harvard University, Cambridge, Mass.
BRIGHAM, DR. CARL C., Department of Psychology, Princeton University, Princeton, New Jersey.
BROYLER, MR. CECIL R., Department of Psychology, Princeton University, Princeton, New Jersey.
BROWN, DR. J. F., University of Kansas, Lawrence, Kansas.
BURKS, MISS BARBARA STODDARD, Institute Rousseau, University of Geneva, Geneva, Switzerland.
BURNHAM, MR. PAUL S., Yale University, New Haven, Connecticut.
BUROS, MR. OSCAR K., Rutgers University, New Brunswick, New Jersey.

CARTER, DR. HAROLD D., 2739 Bancroft Way, Berkeley, California.
CHAMPNEY, MR. HORACE, 246 West Woodruff Avenue, Columbus, Ohio.
CHANT, S. N. F., University of Toronto, Toronto, Canada.
CHESIRE, MISS LEONA, University of Chicago, Chicago, Illinois.
CONRAD, DR. HERBERT S., Department of Education, University of California, Berkeley, California.
COOMBS, MR. CLYDE H., Department of Psychology, University of California, Berkeley, California.
COREY, MR. STEPHEN MAXWELL, Department of Education, University of Nebraska, Lincoln, Nebraska.
COWLES, MR. ALFRED, 3RD, 301 Mining Exchange Bldg., Colorado Springs, Colo.
COX, MISS GERTRUDE, Iowa State College, Ames, Iowa.
CULLER, DR. ELMER K., Department of Psychology, University of Illinois, Urbana, Illinois.
CURETON, DR. EDWARD E., Polytechnic Institute, Auburn, Alabama.

DODD, STUART C., American University of Beirut, Lebanon, Syria.
DUNLAP, DR. JACK W., The Graduate School, Fordham University, New York City.

EDGERTON, DR. HAROLD, Room 1109, U. S. Department of Labor, United States Employment Office, Washington, D. C.
ELLICKSON, MR. J. C., 1916 G. Street, N. W., Washington, D. C.
ENGELHART, DR. MAX D., 5454 S. Greenwood Avenue, Chicago, Illinois.
EURICH, DR. ALVIN C., Department of Education, University of Minnesota, Minneapolis, Minnesota.

FERNBERGER, DR. S. W., College Hall, University of Pennsylvania, Philadelphia, Pennsylvania.
FINDLEY, DR. WARREN G., Cooper Union, Columbia University, New York City.
FLANAGAN, DR. JOHN C., 500 West 116th St., New York City.
FRANZEN, DR. RAYMOND H., 155 East 44th Street, Chicago, Ill.
FREYD, DR. MAX, Social Security Board, Washington, D. C.

GARRETT, DR. HENRY E., Department of Psychology, Columbia University, New York City.
GENGERELLI, DR. J. A., University of California at Los Angeles, Los Angeles, California.
GREENE, DR. E. B., University of Michigan, Ann Arbor, Michigan.
GRIFFIN, DR. HAROLD B., Nebraska State Teachers College, Wayne, Nebraska.
GUILFORD, DR. JOY PAUL, Department of Psychology, University of Nebraska, Lincoln, Nebraska.
GULLIKSEN, DR. HAROLD O., University of Chicago, Chicago, Illinois.
GUTHRIE, DR. E. R., University of Washington, Seattle Washington.

HARSH, MR. CHARLES, Department of Psychology, University of California, Berkeley, California.
HARTMAN, DR. GEORGE W., Teachers College, Columbia University, New York City.
HELSON, DR. HARRY, Bryn Mawr College, Bryn Mawr, Pennsylvania.
HENRY, DR. E. R., University College, New York University, University Heights, New York City.
HINCKLEY, DR. ELMER D., Department of Psychology, University of Florida, Gainesville, Florida.
HO, DR. CHING JU, Lsing Hua University, Peiping, China.
HOLZINGER, DR. KARL, University of Chicago, Chicago, Illinois.
HORST, DR. PAUL A., Procter and Gamble Company, Cincinnati, Ohio.
HOTELLING, DR. HAROLD, Department of Economics, Columbia University, New York City.
HULL, DR. CLARK L., Department of Psychology, Yale University, New Haven, Connecticut.

JENKINS, DR. RICHARD L., Institute for Juvenile Research, 907 South Lincoln Street, Chicago, Illinois.
JOHNSON, DR. PALMER, College of Education, University of Minnesota, Minneapolis, Minnesota.
JONES, DR. LL. WYNN, 7 Bideford Avenue, Leeds 8, England.

KOCH, DR. HELEN LOIS, University of Chicago, Chicago, Ill.
KORNHAUSER, DR. ARTHUR W., University of Chicago, Chicago, Ill.
KUDER, MR. G. FREDERIC, 2344 Neil Avenue, Columbus, Ohio.
KURTZ, DR. ALBERT K., 1170 Elm Park Drive, Cincinnati, Ohio.
KUZNETS, MR. GEORGE, Department of Psychology, University of California, Berkeley, California.

LARSON, MR. S. C., Carleton College, Northfield, Minnesota.
LAZARSFELD, DR. PAUL F., University of Newark, 40 Rector Street, Newark, New Jersey.
LEDGERWOOD, DR. RICHARD, Southeastern Teachers College, Durant, Oklahoma.
LEWIN, DR. KURT, Child Welfare Research Station, Iowa City, Iowa.
LINDQUIST, MR. E. F., Iowa State University, Iowa City, Iowa.
LORGE, DR. IRVING, Department of Psychology, Columbia University, New York City.
LUNDBERG, DR. GEORGE A., Bennington College, Bennington, Vermont.
LURIE, DR. WALTER, 5432 Kenwood Avenue, Chicago, Ill.

MARGINEANU, DR. NICHOLAS, Institute of Psychology, Il Regala, Cluf, Rumania.
MAY, DR. MARK A., Department of Psychology, Yale University, New Haven, Connecticut.
MCGEOCH, DR. JOHN A., Department of Psychology, Wesleyan University, Middleton, Conn.

McNamara, Mr. Walter J., Department of Psychology, University of Minnesota, Minneapolis, Minn.
McNemar, Dr. Quinn, Stanford University, California.
Metfessel, Dr. Milton, University of Southern California, Los Angeles, Calif.
Miner, Dr. James Burt, Department of Psychology, University of Kentucky, Lexington, Kentucky.
Monroe, Dr. W. S., College of Education, University of Illinois, Urbana, Illinois.
Moore, Dr. Thomas Verner, Catholic University of America, Washington, D. C.
Mosier, Mr. Charles, University of Chicago, Chicago, Ill.
Muenzinger, Dr. Karl F., University of Colorado, Boulder, Colorado.
Murchison, Dr. C., Clark University, Worcester, Massachusetts.

Newland, Mr. H. C., Department of Education, Edmonton, Alberta, Canada.
Newman, Mr. Wilson J., Department of Psychology, University of Chicago, Chicago, Ill.

Olson, Prof. Willard C., University of Michigan, Ann Arbor, Michigan.
O'Rourke, Dr. L. J., Director of Research in Personnel Adm., U. S. Civil Service Commission, Washington, D. C.
Otis, Mr. Arthur S., World Book Company, Yonkers-on-Hudson, New York.
Otis, Dr., J. L., 1109 Department of Labor Building, 14th and Constitution Avenue, Washington, D. C.

Pace, Mr. C. Robert, Room 11, Library, University of Minnesota, Minneapolis, Minnesota.
Paterson, Prof. Donald G., Department of Psychology, University of Minnesota, Minneapolis, Minnesota.
Peatman, Dr. John Gray, College of the City of New York, Convent Avenue & 139th Street, New York City.
Peters, Dr. C. C., State College, Pennsylvania.
Piaggio, Mr. H. T. H., University College, University Park, Nottingham, England.
Pieron, H., Sorbonne, Paris, France.

Rashevsky, Dr. N., University of Chicago, Chicago, Ill.
Reitz, Dr. Wilhelm, United States Public Health Service, 153 East Elizabeth St., Detroit, Michigan.
Remmers, Dr. Hermann H., Department of Education, Purdue University, Lafayette, Indiana.
Reymert, Dr. Martin L., Mooseheart Laboratory for Child Research, Mooseheart, Illinois.
Richardson, Dr. Marion W., University of Chicago, Chicago, Ill.
Roff, Dr. Merrill E., University of Indiana, Bloomington, Indiana.
Roley, R. D., Eastern Gas and Fuel Associates, 250 Stuart Street, Boston, Mass.
Royer, Dr. Elmer B., 916 West 4th Avenue, Stillwater, Oklahoma.
Ruch, Dr. Floyd, Department of Psychology, University of Illinois, Urbana, Ill.
Rulon, Dr. Phillip Justin, Department of Education, Harvard University, Cambridge, Mass.
Ruml, Dr. Beardsley, Macy and Company, New York City.
Russell, Dr. James T., University of Chicago, Chicago, Ill.

Saffir, Dr. Milton A., 1520 South Hamlin Avenue, Chicago, Ill.
Schultz, Dr. Henry, University of Chicago, Chicago, Ill.
Shock, Dr. Nathan Wetherill, Department of Psychology, University of California, Berkeley, California.
Sims, Dr. V. M., Box 1021, University, Alabama.
Slesinger, Dr. Donald, National Association of Housing Officials, 730 Jackson Place, Washington, D. C.
Snoddy, Dr. George S., Department of Psychology, Indiana University, Bloomington, Indiana.
Spearman, Dr. C., University College, University of London, London, England.
Stalnaker, John M., University of Chicago, Chicago, Ill.
Stephenson, Dr. W., Psychology Department, University College, Gower Street, London, W. C. 1, England.

STODDARD, MR. GEORGE D., Department of Psychology, University of Iowa, Iowa City, Iowa.
STOUFFER, DR. SAMUEL, University of Chicago, Chicago, Ill.
STOY, DR. EDWARD G., 68 Post Street, San Francisco, California.
STRANG, DR. RUTH, Teachers College, Columbia University, New York City.
STRONG, DR. EDWARD K., JR., Stanford University, California.
STUIT, MR. D. B., Carleton College, Northfield, Minnesota.
SWANN, DR. HOWARD, 5746 Maryland Avenue, Chicago, Illinois.

TERMAN, DR. LEWIS M., Stanford University, California.
THOMSON, DR. GODFREY, University of Edinburgh, Edinburgh, Scotland.
THORNDIKE, DR. E. L., Columbia University, New York City.
THURSTONE, DR. L. L., University of Chicago, Chicago, Ill.
THURSTONE DR. THELMA G., 5642 Kimbark Avenue, Chicago, Ill.
TIEGS, DR. E. W., 1300 Transportation Building, 7th and Los Angeles Streets, Los Angeles, California.
TOOPS, DR. HERBERT A., Department of Psychology, Ohio State University, Columbus, Ohio.
TRAXLER, DR. ARTHUR E., Graduate Education Building, University of Chicago, Chicago, Illinois.
TRYON, DR. ROBERT CHOATE, Department of Psychology, University of California, Berkeley, California.
TUCKER, MR. L. R., 6237 Drexel Ave., Chicago, Ill.
TYLER, DR. RALPH W., Ohio State University, Columbus, Ohio.

UPSHALL, DR. C. C., State Normal School, Bellingham, Washington.

VAN STEENBERG, MR. N., University of Chicago, Chicago, Ill.
VAUGHN, MR. CHARLES L., Wayne County Training School, Northville, Michigan.
VITELESS, DR. MORRIS S .,Department of Psychology, University of Pennsylvania, Philadelphia, Pennsylvania.
VON BONIN, DR. GEHRHARDT, 1853 West Polk Street, Chicago, Illinois.

WALKER, DR. HELEN M., Teachers College, Columbia University, New York City.
WALLACE MR. R. R., 6231 Greenwood Avenue, Chicago, Illinois.
WHITE, DR. RALPH K., Wesleyan University, Middletown, Connecticut.
WILEY, MR. LLEWELLYN N., Department of Psychology, University of Illinois, Urbana, Illinois.
WILKS, DR. F. S., Department of Mathematics, Princeton University, Princeton, New Jersey.
WILLIAMS, DR. ROBERT D., Department of Psychology, Ohio State University, Columbus, Ohio.
WOLFLE, DR. DAEL L., University, Mississippi.
WOOD, DR. BEN D., Columbia University, New York City.
WOODROW, DR. HERBERT, Department of Psychology, University of Illinois, Urbana, Illinois.